

**DESIGN METHODOLOGY TO CHARACTERIZE AND
COMPENSATE FOR PROCESS AND TEMPERATURE
VARIATION IN DIGITAL SYSTEMS**

A Dissertation
Presented to
The Academic Faculty

By

Minki Cho

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
December 2012

Copyright © 2012 by Minki Cho

DESIGN METHODOLOGY TO CHARACTERIZE AND COMPENSATE FOR PROCESS AND TEMPERATURE VARIATION IN DIGITAL SYSTEMS

Approved by:

Professor Saibal Mukhopadhyay, Advisor
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Sudhakar Yalamanchili
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Sung Kyu Lim
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Linda S. Milor
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Satish Kumar
School of Mechanical Engineering
Georgia Institute of Technology

Date Approved: September, 2012

*Dedicated to my parents Hanguk Cho and Heekyung Byun,
my wife Sumi Choi, for their endless love.*

ACKNOWLEDGEMENTS

I wish to thank my advisor, Professor Saibal Mukhopadhyay for guiding and advising me during the course of my Ph.D research. I am sincerely grateful for his support, encouragement, and trust in me. I also would like to thank Professor Sudkar Yalamanchili, Professor Sung Kyu Lim for invaluable suggestions on my research. I also would like to thank Professor Satish Kumar and Professor Linda Milor for improving my dissertation as a committee member. I also wish to thank Dr. Arijit Raychowdhury and Dr. Rahul Rao for practical feedback and suggestions.

I would like to thank GREEN lab members, Jeremy Tolbert, Suhbo Chatterjee, Nikhil Sathe, Kwanyeob Chae, Denny Lie, Boris Alexandrov, Amit Trivedi, Khondkar Ahmed, and Wen Yueh for their valuable suggestions and discussions.

I wish to thank GTCAD lab members, Dr. Daehyun Kim, Chang Liu, Moongon Jung for their valuable comments and feedback. I also would like to thank MiNDS lab member, Man Prakash Gupta for our collaboration.

I would like to thank my friends who shared hard time and gave me an opportunity to get away from it during tough Ph.D student life: Jaeyoung Choi, Jihun Oh, Wanik Cho, Yongnam Cho, Seungyun Choi, Jehoon Lee, Sehun Kim, Soonkyo Jung, and Honggab Kim.

In particular, I am grateful to my wife Sumi Choi and parents Hanguk Cho and Heekyung Byun, who provide endless supports, encouragement, and precious love for my life. I also thank my sister Jungyoon Cho for consistent support, encouragement and understanding.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iv
LIST OF TABLES	ix
LIST OF FIGURES	x
SUMMARY	xv
I INTRODUCTION	1
II ORIGIN AND HISTORY OF THE PROBLEM	4
2.1 Conventional Digital System Design.....	4
2.2 Need for Characterizing and Compensating for Process Variation	4
2.3 Need for Characterizing and Compensating for Temperature Variation.....	5
2.4 Prior Work for Characterizing and Compensating for Process and Temperature Variation	6
2.4.1 Process Variation	6
2.4.2 Temperature Variation	7
III DESIGN METHOD TO COMPENSATE FOR PROCESS VARIATION IN LOW-POWER MULTIMEDIA APPLICATIONS	10
3.1 Introduction.....	10
3.2 Memory Failures and Voltage Scaling	13
3.2.1 Access Failure.....	13
3.2.2 Read-Disturb Failure.....	14
3.2.3 Write Failure	14
3.3 Reconfigurable Accuracy-Aware Static Random Access Memory (SRAM)...	15
3.3.1 Proposed SRAM Architecture	15
3.3.2 Design Considerations	17
3.3.3 Array-Power Estimation	22
3.4 System-Level Image-Quality Estimation.....	24
3.5 Simulation Results	28
3.5.1 Circuits Simulation	29

	3.5.2	System-Level Simulation of Images	29
IV		DESIGN METHOD TO CHARACTERIZE AND COMPENSATE FOR PROCESS VARIATION IN 3D ICS	35
	4.1	Need for Testing and Characterization of Process Variation in 3D ICs	35
	4.2	Challenges of Design for Testing and Characterization Structure of the TSVs	36
	4.3	Electrical Effects of TSV Defects	37
	4.4	Testing and Signal-Recovery Structure for TSV	39
	4.4.1	The Basic Structure and Operating Principle	40
	4.4.2	Detailed Circuit Design of the Test and Recovery Circuit	44
	4.5	Simulation Results	47
	4.5.1	Application to Pre-bond TSV Tests and Signal Recovery	47
	4.5.2	Application to Post-bond TSV Tests and Signal Recovery	51
	4.6	System-Level Full-Chip Analysis	53
	4.6.1	Target 3D Structure and Design Flow	54
	4.6.2	Delay and Power Impact of TSV Test Circuits	55
V		DESIGN METHOD TO CHARACTERIZE AND COMPENSATE FOR TEMPERATURE VARIATION IN MANY-CORE SYSTEMS	57
	5.1	Burn-in Test	57
	5.2	Challenges of Burn-in Test for Many-Core System	59
	5.3	Modeling and Simulation Framework	60
	5.3.1	Power Models for Cores	61
	5.3.2	Thermal Modeling Framework	62
	5.3.3	Modeling of Power Supply Noise	63
	5.4	Adaptive Spatiotemporal Power Migration (ASTPM) Architecture	63
	5.4.1	Spatiotemporal Power Migration (STPM)	63
	5.4.2	Adaptive Spatiotemporal Power Migration	65
	5.4.3	Overall Test Methodology	66
	5.5	Simulation Results	67
	5.5.1	Challenges in Many-Core Burn-in Test	68
	5.5.2	Spatiotemporal Power Migration	70

5.5.3	Need for Adaptive Spatiotemporal Power Migration	73
5.5.4	Adaptive Spatiotemporal Power Migration	74
5.6	Burn-in Time Estimation	75
5.7	Characterization of the Effect of Migration Interval in Normal Operation	77
5.7.1	Thermal Impact of the Migration Interval	77
5.7.2	Electrical Impact of Migration Interval and Analysis of Tradeoffs...	81
5.7.2.1	Effect on Core-to-Core Delay and Leakage Variations...	81
5.7.2.2	The Power Overhead and Overall Power Saving	82
VI	CHARACTERIZATION OF INVERSE TEMPERATURE DEPENDENCE AND DIGITAL THERMAL SENSOR	84
6.1	Test-chip Organization.....	84
6.1.1	Poly-resistor-based heater	85
6.2	Analog Sensor.....	86
6.3	Digital Sensor.....	87
6.3.1	Ring Oscillator (RO)	88
6.3.2	Pulse Generator and Level Converter	89
6.4	Design Challenge in Wide-operating Range and Low VDD: Inverse Temperature Dependence	92
6.4.1	Measurement Results	95
6.4.2	The Impact of Device Type	97
6.4.3	The Impact of Circuit Type.....	97
6.4.4	The Impact of Process Variation.....	98
6.5	Conclusion	100
VII	POST-SILICON THERMAL PREDICTION	101
7.1	Introduction.....	101
7.2	Contributions	103
7.3	Related work and Novelty	105
7.4	Mathematical Approach.....	107
7.4.1	Modeling the MIMO Thermal System with Leakage-temperature Iteration	107
7.4.2	Methods for Thermal System Identification	110

7.5	Applications of TSI to Thermal Modeling of Many-Core Processors.....	110
7.5.1	Baseline Thermal Simulator used for Verification of the Proposed Approach	110
7.5.2	Thermal System Identification for Many-Core Processors	111
7.5.3	Accuracy of TSI-based Thermal Prediction	114
7.5.4	Application to Post-Silicon Thermal Prediction	116
7.6	Validation through Test-chip	118
7.6.1	Implementation	119
7.6.2	Measurement Results	120
7.7	Future Work	130
7.8	Conclusion	131
VIII	CONCLUSION	132
8.1	Summary and Contribution.....	132
8.2	Future Research	134
	APPENDIX A.....	136
	REFERENCES	139

LIST OF TABLES

1	Assignment of the control signals for the test circuit	45
2	Power analysis: power overhead of migration.....	83
3	Summary of measurement conditions.....	86
4	Impact of type of devices on temperature dependence	97
5	Average error for 3 power patterns	128

LIST OF FIGURES

1	Basic concept of accuracy-aware low-power array	11
2	Effect of different error rates of LOBs and HOBs.....	12
3	Effect of supply voltage on SRAM cell failures	14
4	Proposed accuracy-aware low-power array	16
5	Effect of PMOS size of the ReconfigInv on WL delay. D is WL delay through the reconfiguration network, and D_{org} is the original WL delay	19
6	Layout of one row of the proposed reconfigurable SRAM showing the reconfiguring inverters.....	19
7	Simulation framework	25
8	Effect of supply voltage scaling on (a) image quality and (b) power saving	30
9	Image with Vdd scaling (a) all bit at 0.4V and (b) $L_{bit} = 4$ and $V_{low} = 0.4V$	31
10	Reconfiguration: (a) image quality and power tradeoff for a given V_{low} , 3D plots showing the co-reconfiguration of L_{bit} and V_{low} and its effect on (b) image quality and (c) power saving	32
11	Quality degradation of different images considering multiple MC runs	33
12	Impact of transient noise on repeated reading of same image (MSSIM normalized to that for 1 st read).....	34
13	Schematic illustration of TSV shorts at the pre-bond stage (top) and variation in the resistance at the post-bond stage (bottom).....	36
14	Effect of TSV short: (a) driver-receiver combination, (b) signal swing, (c) delay, and (d) driver and receiver power.....	38
15	Post-bond TSV resistance: (a) driver-receiver, (b) effect on signal swing, (c) signal slew, and (d) signal delay.....	39
16	The basic test structure and characterization policy	41
17	Test structure for (a) pre-bond test of output TSV, (b) post-bond characterization, and (c) the logic diagram of the overall test structure	42
18	Circuit schematic of the proposed test circuit.....	45
19	The overall test philosophy	46

20	Waveform of operation: (a) detection input, (b) recovery input, (c) detection output, and (d) recovery output.....	48
21	Statistical simulation results showing the functionality during the pre-bond test	48
22	Sources of detection inaccuracy obtained using Monte-Carlo simulations in 90nm CMOS: (a) V_{TSV} variation, (b) detection error, (c) effect of offset variation, and (d) driver-TTI mismatch	50
23	(a) Detection of TSV groups and (b) effect of signal recovery on pre-bond yield	51
24	Post-bond testing and recovery: (a) detection with the test circuit, (b) statistical simulation results showing the functionality at post-bond test	52
25	Post-bond testing and recovery: (a) effect on signal swing, (b) signal slew, and (c) delay.....	53
26	The 3D system-level analysis: (a) the target 3D structure and TSV sizes and (b) full-chip layouts of the designed 3D stack for FFT256_8 circuit	54
27	The 3D design flow used in this work	55
28	The area overhead of the proposed structure from system-level analysis	56
29	Burn-in basics: (a) bathtub curve and (b) effect of temperature on acceleration.....	58
30	Modeling framework	61
31	(a) Concepts of spatiotemporal power migration and (b) spatial and temporal difference	64
32	Adaptive spatiotemporal power migration.....	67
33	Effect of temperature on leakage with process variations	68
34	Self-consistent leakage-temperature simulations.....	69
35	Burn-in challenges for many-core: (a) Ldi/dt droop, (b) supply settling time, and (c) non-uniformity in the thermal profile	70
36	Temperature rise with and without STPM.....	71
37	The effect of migration interval on (a) max temp. vs. time, (b) max temp. vs. number of active cores at 6 sec, (c) spatial difference vs. time interval, and (d) temporal difference vs. time interval	72
38	The effect of number of active cores on (a) max temperature and (b) spatial difference	72
39	The impact of process variation on (a) leakage power distribution, (b) max temperature, and (c) spatial difference.....	73

40	Effect of ASTPM on (a) number of active cores, (b) core-to-core temperature variations, and (c) core-to-core temporal difference variations.....	75
41	Effect of ASTPM on burn-in time on (a) acceleration factor for different corners, (b) normalized burn-in time for different corners, (c) relation between NSSC and leakage current, and (d) burn-in time improvement with leakage binning	77
42	Thermal behavior of many-core system	78
43	Effect of time-slice on random migration: (a) max temperature (b) core-to-core temperature distribution, (c) spatial difference, and (d) temporal difference. The 100K refers to 100,000 clock cycles i.e. 33 μ s for a 3GHz clock	79
44	Effect of time-slice on: (a) max temperature, (b) max spatial difference, (c) max temporal difference , (d) time-slice for a target max temperature on ratio, (e) time-slice for a target spatial difference in different on ratio, and (f) time-slice for a target temporal difference in different on ratio	80
45	Chip temperature with run-time variations in the number of active cores	81
46	Effect of power migration on (a) core-to-core delay distribution, (b) core-to-core leakage distribution, (c) chip delay (i.e., maximum core delay), and (d) chip leakage power	82
47	The die photo showing the organization of the test chip	85
48	Poly-resistor-based heater	86
49	Analog thermal sensor	87
50	Measured results showing the operation of the temperature sensor (T1) and the heater	87
51	Digital sensor structure: (a) Overall block diagram and (b) detail of the ring oscillators	88
52	Pulse generator.....	89
53	Correlation between Power and Frequency of Digital Sensor: (a) correlation between power and frequency with nominal VDD, (b) correlation between power and normalized frequency with nominal VDD, (c) correlation between power and frequency with low VDD, and (d) correlation between power and normalized frequency with low VDD.	91
54	Correlation between Temperature and Frequency of Digital Sensor: (a) correlation between temperature and frequency with nominal VDD, (b) correlation between temperature and normalized frequency with nominal VDD, (c) correlation between temperature and frequency	

	with low VDD, and (d) correlation between temperature and normalized frequency with low VDD.....	92
55	Inverter chain RO waveform at different V_{DD} and temperature	96
56	Measured results showing temperature dependence of an inverter chain RO	96
57	Impact of the LVT and HVT devices on ZTC	97
58	Measured ZTC points for different circuit paths considering process variations. The data from 20 chips are shown	99
59	Measured correlation between ZTC points and frequency for the inverter chain based RO	100
60	Illustration of the need for post-silicon transient thermal analysis considering process variation: (a) the interaction of a (average for all input condition) and temperature in a NAND2 gate considering different process corners (HVT – High threshold voltage, NVT – nominal threshold voltage, and LVT – low threshold voltage corner). (b) the effect of such interaction for an example self-consistent thermal simulation (using distributed RC network) considering a square wave dynamic power profile (e.g. turning on and off a the chip after a time-interval) and leakage of 10million NAND2 gate	102
61	Overall methodology of post-silicon prediction of the transient thermal field. The method uses the time-frequency duality to extract thermal system in frequency domain using post-silicon measurement and use that to predict transient temperature profile.....	104
62	Mathematical principle of the proposed approach. The thermal system is considered as a MIMO system.....	108
63	(a) Transient power traces of exemplary benchmarks for SPEC2006 applications (b) The frequency response of the power traces	111
64	Core gating based approach to power spectra generation. (a) Sleep transistor signal (5 MHz) and (b) power pattern (5 MHz)	112
65	The thermal filter extraction through small signal simulation (ideal approach for filter extraction) and sleep control based power/thermal measurement.....	113
66	Filter behavior of thermal system: distance between source core and observation node	114
67	Estimation error in transient variation of temperature for a typical core in the 64 core system. The simulations were performed considering random workloads created for all 64 cores using random assignments of benchmark applications for SPEC2006 suites	115

68	Estimation error in instances of the spatial thermal field at different time points.....	115
69	Accuracy of the proposed approach considering random workload.....	116
70	The application of TSI based approach on the prediction of impact of process variation on transient temperature: (a) the effect of leakage-temperature interaction and (b) time-domain temperature variation	117
71	The application of TSI based approach on the prediction of impact of the conductivity of thermal stack (TIM, spreader, and heat sink) variation on transient temperature: (a) the effect of thermal conductivity in the extracted filter (b) time-domain temperature variation.....	118
72	Post-silicon thermal characterization and prediction with chip-to-chip variation in leakage and thermal conductivity	118
73	The die photography	119
74	Experiment hardware setup: (a) overall schematic (b) test structure photo.....	120
75	Oscilloscope waveform of thermal sensor and power	121
76	Thermal system extraction.....	122
77	Extracted thermal filters: (a) 3 different locations (b) thermal filter variation for 10 chips (Sensor 1) (c) thermal filter variation for 10 chips (Sensor 2), and (d) thermal filter variation for 10 chips (Sensor 3)	123
78	Hardware measurement: input power patterns	124
79	Temperature prediction: (a) for power pattern 1 (b) for power pattern 2	125
80	Temperature prediction: (a) for Sensor with pattern 1, (b) with pattern 2, (c) for sensor 2 with pattern 1, (d) with pattern 2, € for sensor 3 with pattern 1, and (f) with pattern 2.....	127
81	The application of a digital sensor	129
82	Phase characteristic of thermal systems.....	130
83	Estimation error for a representative core for a filter-based implementation.....	131

SUMMARY

Over the years, the number of transistors in the integrated circuits has increased following Moore's law. Device scaling and new process technologies have allowed Moore's law to continue down to the nanometer regimes. However, in the nanometer regimes, the increased process variation results in the large variation in the performance, which translates into the yield loss of a digital system. A tight power budget worsens tolerance to noise and reliability of the systems. In addition to the device scaling, new process technologies such as a through-silicon-via (TSV) technology to improve the performance of digital systems have emerged. However, since new technology is not matured at the early stage, digital systems that use the TSVs may suffer from the reliability issues for quite amount of time.

The continuous increase in the total chip power and power density in successive generation increases chip temperature. Moreover, the increasing variation in the manufacturing process adds to the variations in the leakage power and hence further increases the chip power. A higher chip temperature degrades circuit performance, increases leakage power, degrades circuit reliability, and reduces cooling efficiency. The increasing chip leakage coupled with the increase in the number of digital blocks in the systems can lead to unsustainable increase in chip power and die temperature imposing stringent challenges in the test and normal operation of digital systems.

The main objective of this dissertation is to investigate a design methodology that can characterize and compensate for process and temperature variation. First, a design methodology is discussed to handle process variation in low-power memory for image processing application.

This is followed by a design technique to characterize and recover TSV-defect-induced signal degradation in a 3D integrated circuit. For thermal variation, the challenge of thermal sensor design is presented, and the characteristics of analog and digital thermal sensors are analyzed through a test-chip. The inverse temperature dependence in the digital logics is characterized through hardware to help better thermal sensor design in wide operating voltage design. The spatiotemporal power migration is proposed as a methodology to handle thermal issues in digital systems both during the test and normal operation. The power migration continuously distributes the generated heat in space and time to control chip temperature. To enable this approach a unique method is developed, and verified through hardware for post-fabrication characterization of thermal system and prediction of transient variation in chip temperature.

CHAPTER I

INTRODUCTION

The semiconductor industry continues to scale down silicon devices to achieve better performance of digital systems for a given power budget. Scaling down the devices provides a lower delay of systems, higher operating frequency, lower power consumption, and lower manufacturing costs. However, as the feature size reaches nanometer nodes, the increasing process variation results in the large variation in performance and leakage of a digital system. For low-power applications, because of a tight power budget, tolerance to noise decreases significantly with process variation, and reliability issues become critical. Digital systems also suffer from variations in a new process. For example, three-dimensional integrated circuit (3D IC) is a promising technology for better performance and low power consumption because of higher bandwidth and shorter interconnects compared to conventional 2D ICs. A new technology through-silicon-via (TSV) is necessary to build a 3D IC. However, since the TSV process is not perfect one, the defect and variability of the TSV affect the overall yield and the performance of 3D ICs. Therefore, characterizing and compensating for process variation are important to improve the yield and the reliability of digital systems. The proposed research presents a design method to deal with process variation in a low-power multimedia application and a design method to characterize and compensate for process variation in 3D ICs.

In addition to the process variation, the dynamic temperature variation can affect the performance of a system. In digital systems, the chip temperature determines the operating frequency of systems. To operate the digital systems at the target frequency, the temperature

needs to be maintained below the threshold temperature. When the chip temperature increases above the threshold, a system can reduce the supply voltage or increase the fan speed to reduce the chip temperature. However, reducing the supply voltage decreases the throughput of the systems, and applying additional cooling power increases the total power. Therefore, a method that can maintain both the throughput and the temperature needs to be considered. The proposed research includes a design method that can redistribute the generated heat within the chip over space and time to reduce the maximum temperature and to maintain a uniform thermal map. This proposed method is applied to both burn-in test and normal operation.

Various design methodologies including the proposed approach for improving the thermal field make a decision based on the transient variation in the chip temperature, which is determined by physical parameters (e.g., thermal resistance and thermal capacitance). The physical parameters of a chip are different depending on the material and can change over time. Therefore, to manage the thermal field more efficiently, characterizing the thermal physical parameters is necessary. The proposed research includes the characterization of the thermal field through a thermal-filter-based approach.

This dissertation is organized as follows. Chapter 2 reviews the overview of the problems in conventional digital system design and the importance of characterizing and compensating for process/temperature variation. Chapter 3 presents a design method to compensate for process variation in low-power multimedia applications. Chapter 4 discusses a design method to characterize and compensate for process variation in 3D ICs. Chapter 5 presents an operational method to characterize and compensate for temperature variation in many-core systems both during the test and normal operation. Chapter 6 discusses thermal sensor design challenge and inverse temperature dependence. Chapter 7 discusses a design methodology for thermal system

identification and transient temperature prediction. This research is developed through both simulation and hardware measurement. Chapter 8 summarizes the dissertation and suggests the future research.

CHAPTER II

ORIGIN AND HISTORY OF THE PROBLEM

2.1 Conventional Digital System Design

As the physical feature size of the transistors reduces, the impact of variations in process on the systems increases. To compensate for the effect of process and temperature variation, system designers estimate a process corner and a “possible” worst thermal profile in a space and time. Based on the estimated worst-case scenarios, designers add additional design margin (i.e., over-design) to compensate for the effects of process and temperature variation. Worst-case corner is estimated and predicted through multiple Monte-Carlo simulations. However, adding sufficient design margin can increase the power and area overhead. On the other hand, other designers have explored to add minimum design margin to pass the functionality under an estimated “possible” minimal variations, which can put digital systems at risk. Depending on the process corner and thermal condition, the system designed by a conventional design method can operate fast unnecessarily because of the overestimated design margin or fail to function properly because of the underestimated design margin.

2.2 Need for Characterizing and Compensating for Process Variation

Variation in the process results in a significant spread in the performance, leakage, and power consumption. Large variation in performance and leakage causes yield loss and low reliability of digital systems. With the increasing demand for mobile and low-power devices, operating voltage continuously has been reduced. Although scaling the operating voltage of the digital

systems can reduce power dissipation, a low supply voltage increases the sensitivity of circuits to manufacturing variations and increases parametric failures in systems (i.e., a digital system, which is functional at the nominal supply voltage, can fail at a low voltage). From a system perspective, this process variation leads to a high failure rate with voltage scaling and limits the opportunity for power saving.

Along with device scaling, the introduction of a new process such as a 3D technology improves performance with low power consumption, but the process needs significant time to be matured. The through-silicon vias (TSVs) in a 3D stack are the channels for transferring signals between different tiers in a 3D stack. The functionality of a 3D IC strongly depends on the fidelity of signals through TSVs [18, 19]. As the TSV process is not a perfect one, defects can be created while forming the TSVs or while bonding different dies together [18-29]. The defect of the TSV process can determine the performance and yield of 3D ICs. Since maintaining the signal fidelity through TSVs is a primary requirement for 3D system integration, the characterization of the electrical impact of the TSV defects/variations is critical. Therefore, to obtain the better yield and reliability improvement as well as high performance, characterization and compensation for process variation are necessary.

2.3 Need for Characterizing and Compensating for Temperature Variation

The dynamic and leakage power of a chip increases in successive technology generations. Additionally, because of small size of a chip, power density continues to increase, which in turn increases a chip temperature. A high on-chip temperature degrades circuit performance, increases system-leakage power, degrades circuit reliability, and reduces cooling efficiency (i.e., increases cooling energy) [51-62].

The increasing variations in the manufacturing process will result in the variation in leakage power leading to wide chip-to-chip leakage power variations. The variation in dynamic and leakage power causes the fluctuation of the temperature in time and space. Temperature variation across the chip affects the reliability and performance of systems. The variation in temperature across the communicating block can cause performance mismatches, which can result in functional failures in digital systems.

In addition to normal operation, temperature variation during a test affects the quality of test and test time. Specifically, in burn-in test, as the testing conditions include high voltage and high temperature stress, managing temperature becomes difficult and critical. Temperature profile across the chip needs to be uniform and stable during the test to obtain high testing coverage and reduce the burn-in test time. Therefore, the characterization and the compensation for temperature variation need to be considered carefully in digital systems.

Digital circuits supporting wide V_{DD} operating range –from V_{MAX} used in high performance modes down to near threshold voltages (or NTV) used in low power modes – are emerging as key component for next generation DVFS processors and system-on-chips. However, designing across wide V_{DD} range while ensuring correct operation under process (within-die and die-to-die) and temperature variations, is a major challenge. Temperature and voltage variations are further intertwined as temperature increase can have a positive or a negative impact on circuit delay depending on V_{DD} . Thus, it is critical to understand thermal sensor design challenge and temperature dependence for higher performance at low operating voltage as well as a more accurate thermal management.

2.4 Prior Work for Characterizing and Compensating for Process and Temperature Variation

2.4.1 Process Variation

Many design methodologies have been studied to characterize and mitigate the impact of process variation. Adaptive body biasing [46, 56] has been studied in various applications to deal with process variation over the years. Adaptive body biasing applies reverse body bias to increase the threshold voltage and forward body bias to decrease the threshold voltage. This method provides better product yield and the uniform chip-to-chip distribution of performance. Another approach to mitigate the impact of process variation is adaptive supply voltage methods. In this approach, supply voltage varies adaptively to improve the product yield and performance of the systems.

However, in a low-power multimedia application, performance-optimized and reconfigurable systems can be more useful in terms of power saving and the quality of application. Kurdahi et al. [9] and Djahromi et al. [9, 10] have shown that aggressive voltage scaling can be performed in SRAM arrays for multimedia and communication applications exploiting this inherent error tolerance. George et al. [11] and Cheemalavagu et al. [12] have explored the power advantage of using different voltages across a word for arithmetic circuits. Yi et al. [13] have shown that such a concept can be very useful for improving effective yield of video memories.

As the 3D technology gets huge attention from the semiconductor industries, the characterization of the impact of a new process and the compensation for the impact become very critical. Since the functionality of a 3D IC depends on the TSV, the TSV needs to be tested before and after bonding in a 3D IC. While the post-bond testing is necessary to ensure that the customers get defect-free product, the pre-bond testing is more effective to improve the yield. The pre-bond test for detecting resistive shorts and opens has received limited attention in recent literatures [26-29]. Tsai et al. have discussed detection of pin-holes defects in TSVs by using circuits to measure TSV resistance [26]. Chen et al. have proposed to measure the RC time

constant of TSVs and compare that measurement with the known values to determine the existence of faults [27, 28]. However, there is no prior work that presents an on-chip circuit that can simultaneously detect weak/strong defects in TSVs during pre-bond and post-bond test and perform signal recovery under non-catastrophic failures in TSVs.

2.4.2 Temperature Variation

Various architecture-level policies and design methods dealing with temperature variation have been developed over the years. First, to reduce a peak temperature, a method to vary cooling power (i.e., vary a cooling-fan speed) has been applied to high-performance microprocessors. For instance, Intel microprocessors change a fan speed adaptively in runtime, which allows processors to remain within acceptable thermal specifications [70]. The cooling fan operates at a low speed, while a peak temperature is low. If a peak temperature increases beyond the lower-set point, the fan speed will increase until the higher-set point is reached. Second, a method to vary the supply voltage or frequency of the digital systems has been studied. Dynamic voltage scaling (DVS) and dynamic frequency scaling (DFS) have been proposed to handle temperature variation [53, 57]. However, DFS or DVS technique reduces the peak temperature by reducing the heat generation at the expense of reduced performance. Finally, a method to distribute the generated heat across the chip has been researched. Chaparro et al. considered thread migration (TM or core hopping) in multi-core architecture with 16 cores [54]. Donald et al. exploit the distributed nature of multi-core processors and classify the techniques for multi-core thermal management [62]. Ge et al. present a distributed thermal balancing policy to stabilize the operating temperature and improve the performance of many-core systems by distributing task migration with a lightweight agent based on steady-state temperature [63]. In [67], Heat-and-Run performs SMT-thread assignment and CMP-thread migration to maintain throughput under

thermal constraints. Yeo et al. present predictive dynamic thermal management with an advanced future temperature prediction model for multi-core systems [64].

Most of the thermal management research has focused on the peak-temperature reduction. However, significant spatiotemporal non-uniformity in the thermal field is detrimental to both reliability and cooling efficiency. Large and fast temporal variation in temperature (i.e., faster thermal cycles) is also detrimental to circuit reliability. Spatial variation in temperature results in non-uniform delays for different cores leading to reduce performance and functional failures. In addition to normal operation, compensation for temperature variation in a test is critical. Specifically, many studies have been performed about the burn-in test. Since the burn-in-test cost is a significant fraction of the overall budget, the test cost and yield loss during burn-in test needs to decrease [40]. However, the test time, test cost, and yield loss reduction need to be achieved without sacrificing the test coverage. Significant efforts have been directed to reduce the burn-in time with reasonable fault coverage for single or few core processors [41-43].

CHAPTER III

DESIGN METHOD TO COMPENSATE FOR PROCESS VARIATION IN LOW-POWER MULTIMEDIA APPLICATIONS

3.1 Introduction

With the continuous growth of consumer electronics, mobile devices need to perform complex computing as well as diverse sets of images and video/audio processing applications on a very small energy budget, while maintaining the required performance and quality of service [1]. Various studies have shown that memory costs can contribute significantly to the overall power consumption in multimedia systems [2-4]. The optimizations of the number of memory accesses and the reduction of the bus power consumption have been explored as architecture-level methods to reduce memory power [2-5]. High power reduction can be achieved by combining these approaches with the reduction of array power in each access. Scaling the operating voltage of the memory array can reduce the power dissipated in each memory access. However, a low supply voltage increases the sensitivity of circuits to manufacturing variations, which is a particularly serious problem for static random access memories (SRAM). Manufacturing and intrinsic device variations result in a large number of parametric failures in an SRAM array, and the probability of failures increases with voltage scaling [6-8]. An SRAM cell, which is functional at the nominal supply voltage, can fail at a low voltage. From a system perspective, such a failure can lead to a high bit-error rate with voltage scaling and limit power saving, which is a serious problem for SRAM arrays used for data processing. However, even with a non-negligible amount of bit-error rate, image processing and multimedia applications can provide the acceptable quality of service because low-order bits (LOBs) of an image pixel (and many

other signals) are more tolerant to noise than high-order bits (HOBs) in typical multimedia applications. Hence, spatially non-uniform voltage scaling, in which low-order bits (LOBs) operate at a low voltage and high-order bits (HOBs) operate at a high voltage, can achieve appreciable power saving with minimal image or signal quality degradation. Through a system-level experimental analysis, this chapter shows that a high voltage for HOBs and a low voltage for LOBs can improve the yield of SRAM for multimedia applications. The conceptual view of an SRAM architecture with spatial voltage scaling is shown in Fig. 1.

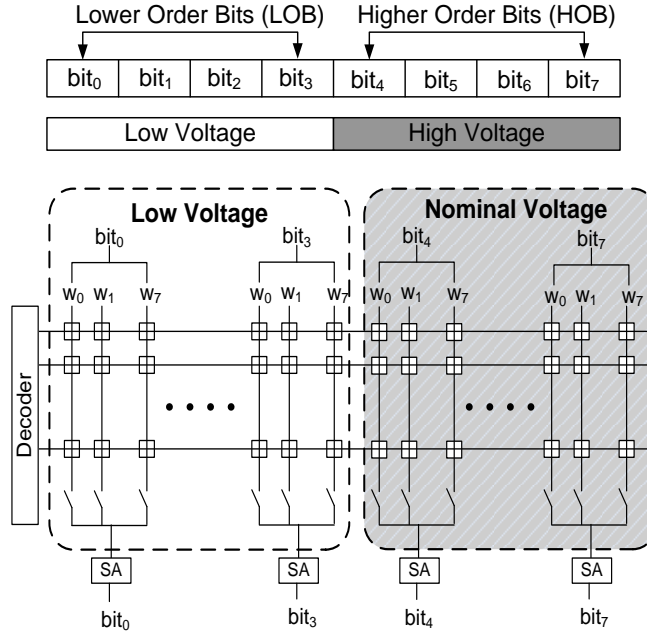


Fig. 1: Basic concept of accuracy-aware low-power array.

Fig. 2 shows the effects of voltage scaling on image quality, which takes into account similar and different failure probabilities (i.e., bit-error rate) for LOBs and HOBs. High failure probability for a memory cell storing the LOBs of image pixels implies a low operating voltage for those cells. In this chapter, the mean structural similarity (MSSIM) index proposed in [15] is used for quality comparison. If two images are visually identical, the MSSIM is 1, and if the image quality is degraded, the MSSIM is low. The MSSIM captures the structural or visual

similarity between two images.



Fig. 2: Effect of different error rates of LOBs and HOBs.

The primary challenge in a mobile system is that the SRAM array is shared among different applications. While image processing applications have inherent error tolerance, data centric applications have less error tolerance. Hence, depending on the error requirements of the application or the characteristics of the image, the array needs to be reconfigured to low-error or high-error modes, which implies a dynamic run-time reconfiguration of the number of bits in the low-voltage mode. The design-time choice (or post-silicon tuning based on the extent of the process variation) of the number of LOBs in the low-voltage (i.e., high-error) mode can lead to varying and unacceptable quality degradation for different images and applications.

As the various failure mechanisms in an SRAM cell depend on the supply, wordline, and bitline voltages of a cell, the run-time reconfiguration requires an innovative architecture. Proper configurations for all the voltage levels are necessary to dynamically modify the number of bits in the low-error and the high-error modes, which cannot be achieved in the standard SRAM VLSI μ -architecture. In other words, spatial voltage scaling in SRAM for multimedia application is not feasible without a reconfigurable SRAM architecture and associated circuit techniques. Note that uniform voltage scaling can achieve such a reconfiguration in the standard SRAM array, but it degrades image quality more aggressively, as shown in Fig. 2. The fine-grid

variation in the supply voltage also increases the complexity of a design.

3.2 *Memory Failures and Voltage Scaling*

Parametric failures induced by random variations in the SRAM cell can be the result of access, read-disturb, and write failures [6-8]. All types of failures increase with a reduction in cell supply, wordline, and bitline voltages (collectively referred to as the “array voltage”). In this analysis, a constant peripheral (i.e., decoder logic and sensing circuits) voltage is considered.

3.2.1 *Access Failure*

If the bit-differential (Δ_{BIT}) developed by a cell at the time of sense-amplifier firing is less than the offset voltage (Δ_{offset}) of the sense amplifiers, access failures occur. The probability of a failure is estimated by access margin ($AM = \Delta_{BIT} - \Delta_{offset}$) as [6]

$$P_{AF} = \Pr\left\{\left(\Delta_{BIT} - \Delta_{offset}\right) \leq 0\right\} \quad (1)$$

Because of a low cell-read current, a low voltage results in a low Δ_{BIT} . Furthermore, a variation in Δ_{BIT} also increases at a low voltage, resulting in a high access failure as shown in Fig. 3. As explained by Wicht et al. [14], the failure is partially compensated by the low offset variations in sense-amplifiers at a low bit-line voltage level (considering that the sense-amplifier supply remains high), which was verified by SPICE Monte-Carlo (MC) simulations in predictive 70nm technology. For a medium frequency of operation mostly used in the multimedia applications (<500MHz), the access failures were observed to be significant only at very low voltage levels (<0.6V). Note that by definition, an access failure is related to the access delay of an array. For a constant frequency of operation, a high access failure at a low voltage essentially captures the effect of the array voltage on the memory delay.

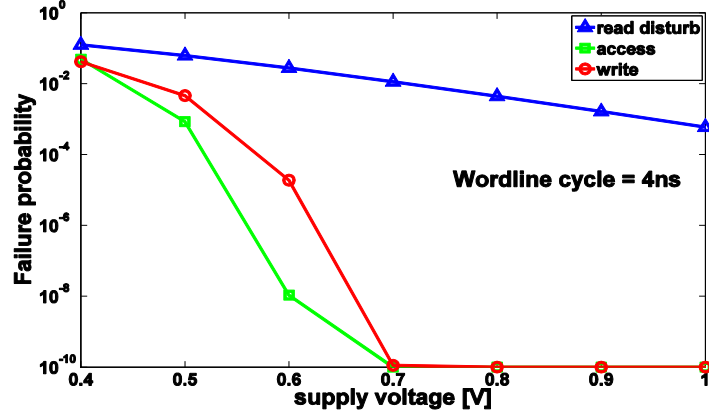


Fig. 3: Effect of supply voltage on SRAM cell failures.

3.2.2 Read-Disturb Failure

Read-disturb failures occur, if the cell data flips while reading a cell. A first-order definition of the read margin (RM) of a cell is the difference between the trip point of the cell inverters (V_{trip}) and the increase in the voltage at the node storing “0” (read-disturb voltage, V_{read}) [6]. The failure event and probability are defined by

$$P_{RF} = \Pr \{ V_{trip} - V_{read} \leq 30mV \} = \Pr \{ RM \leq 30mV \} \quad (2)$$

The disturb failure has weak frequency dependence and dominates the total failure at the nominal array voltage at a medium operating frequency range (<500MHz) as shown in Fig. 3. A low V_{trip} and high variation in V_{READ} increase the disturb failure at a low array voltage.

3.2.3 Write Failure

If a cell cannot be successfully written, write failures occur, which are estimated by evaluating the variation in the write time (i.e., the difference between the wordline edge and the time when the two nodes of the cell achieve the same value) [6, 7]. The write failure event and its probability can be estimated as

$$P_{WF} = \Pr \{ T_{write} > T_{wl} \} = \Pr \left\{ \frac{1}{T_{write}} \leq \frac{1}{T_{wl}} \right\}, \quad (3)$$

where T_{wl} is the time period for which the wordline remains high. As a result of the highly skewed nature of the T_{write} distribution [6, 7], Agawal et al. show that the write failure probability can be better estimated by taking the variation in $(1/T_{write})$ into account [7]. The low strength of a cell-access transistor results in a high write failure at a low array voltage. In addition, at a medium operating frequency range, a write failure is negligible at a nominal supply and becomes significant at a low array voltage, as shown in Fig. 3.

3.3 Reconfigurable Accuracy-Aware Static Random Access Memory (SRAM)

This section presents the proposed reconfigurable accuracy-aware low-power SRAM architecture for image storage. The energy-accuracy tradeoff is performed by applying a low voltage for the cells storing LOBs, while nominal voltage will be applied to the cells storing HOBs as shown in Fig. 1. We propose a reconfigurable architecture which can dynamically modify the number of bits in the low-voltage (L_{bit}) and high-voltage domain. Real-time modification of L_{bit} , depending on the error tolerance of an application, can result in better energy-accuracy tradeoff. A reconfigurable unit is defined as a set of bits connected to the same voltage domain. The reconfiguration can be performed only in reconfigurable units. The number of bits in each reconfigurable unit is defined as the reconfiguration length (RL). The efficiency of the reconfigurable solution depends on the reconfiguration length (RL). The architecture considering the 1-bit reconfigurable unit (bit-by-bit reconfiguration, $RL=1$) is presented. Next, the design considerations for the RL are discussed. For non-image storing applications, all the bits can be configured to the high-voltage domain to reduce errors.

3.3.1 Proposed SRAM Architecture

The overall architecture of the accuracy-aware reconfigurable SRAM architecture is presented in Fig. 4. The conventional array architecture with column multiplexing is considered. In the

column-multiplexing-based architecture, particular bits of different words are grouped together (hereafter referred to as a MUX group). A single read/write circuit is used for a MUX group as shown in Fig. 4. Therefore, entire MUX groups need to be at the same voltage level. The key requirement of the proposed approach is that the cell supply, bitline pre-charge, and write voltage (i.e., the voltage applied to bitline of logic “1”) need to be at the same voltage level. To achieve this requirement, the following modifications are proposed to SRAM array design.

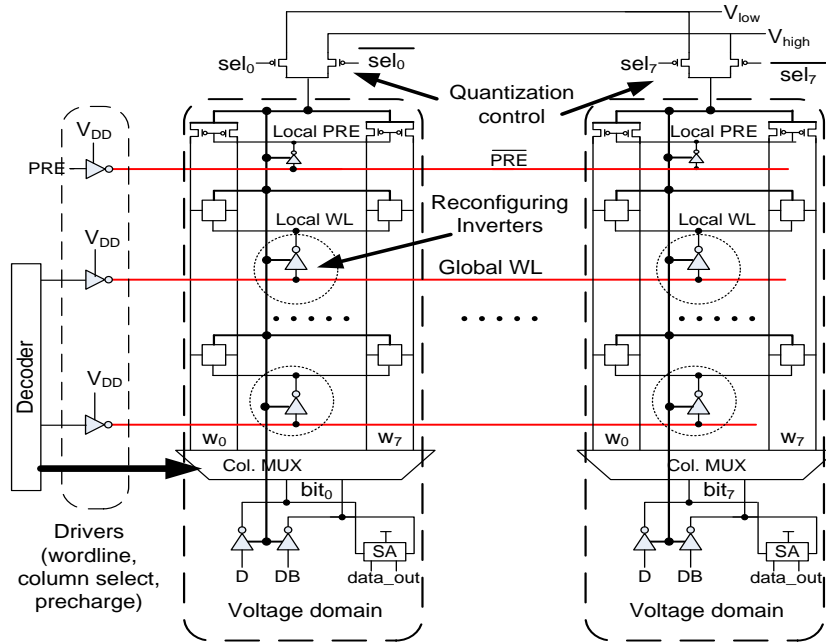


Fig. 4: Proposed accuracy-aware low-power array.

- A column-based supply network for cells will be used as discussed in [16]. The supply networks of all the cells in a MUX group [i.e., number of rows (N_{row}) x number of columns (N_{col}) in a MUX group (N_{mux})] are connected together. The supply networks for different MUX groups are disconnected to allow bit-by-bit reconfiguration.
- The cell supply of a MUX group is connected to the corresponding pre-charge device and the write driver supply. Therefore, changing the pre-charge voltage using a voltage-switching network is sufficient to change the cell supply, bitline, and write voltage for the array.

- The major challenge is to reconfigure the wordline (WL) signal. In a regular array, the WL signal for all the bits is connected. To address this challenge, a reconfigurable wordline structure has been developed as shown in Fig. 4. In this structure, local WLs (LWLs) of a row of cells in a MUX group are connected together. The LWLs of different MUX groups are disconnected. The LWLs are connected to the output of the reconfiguring inverters (ReconfigInvs) which have the global WL (GWL) as the input. The supply of the ReconfigInvs is connected to the supply voltage of that MUX group. For a selected row, GWL is “0”, which makes the LWL same voltage (= “1”) as the cell supply and bitline voltage. The driver of GWL operates at the nominal voltage to eliminate the short-circuit current through the ReconfigInvs in the nominal voltage units. The ReconfigInvs can be applied to also pre-charge and column-selection network for power saving (not necessary for failure reduction).
- The sense-amplifiers driving the output buses need to operate at a high voltage. Therefore, the supply voltage of the sense-amplifiers is not modified. This architecture provides an inherent level conversion for the low-voltage signal from the LOB cells.

3.3.2 Design Considerations

Although the basic structure of the regular array remains unchanged, following elements need careful design considerations.

Reconfiguring Inverters

A ReconfigInv is shared by a MUX group and needs to be designed properly. The low-to-high transition of the local WL is the critical output transition for the ReconfigInv. Therefore, the critical transistor is the PMOS, which needs to be sized to minimize the wordline delay for the low-voltage domain. The total WL delay with ReconfigInvs is given by

$$T \propto \underbrace{\frac{N_{col}C_{bitWL} + N_{bit}(uC_{ax} + C_N)}{sI_{ax}}}_{\text{Driver Delay}} V_{high} + \underbrace{\frac{N_{mux}C_{bitWL} + 2N_{mux}C_{ax}}{u(\alpha I_{ax}/r)}}_{\text{ReconfigInv Delay}} V \quad (5)$$

where $u = W_P(\text{ReconfigInvs})/W_{AX}$, C_N is the capacitance of the NFET ReconfigInvs, r is the NMOS to PMOS current ratio, $s = W_N(\text{driver})/W_{AX}$, C_{bitWL} is the wordline capacitance per cell, C_{ax} is the gate capacitance of the access device, I_{ax} is the effective current of the access device, W_P is the width of the PMOS device in the reconfiguration inverter, W_N is the width of the NMOS device in the reconfiguration inverter, and W_{AX} is the width of the access transistor in SRAM cell width. The factor α represents the reduction of the PMOS current at different voltage domain (i.e., $V = V_{low}$ or V_{high}). The total delay can be minimized as

$$\frac{\partial T}{\partial u} = 0 \Rightarrow u_{opt} = \sqrt{\frac{srN_{mux}(2 + C_{bitWL}/C_{ax})}{\alpha N_{bit}}} \frac{V}{V_{high}} \quad (6)$$

A large value of u_{opt} will be required for the inverters in the low-voltage domain. However, increasing the PMOS size increases the power dissipation in the GWL driver and the area overhead of the array. Therefore, an engineering choice of PMOS size is required. Fig. 5 shows the impact of PMOS transistor size of the ReconfigInv on the WL delay. The ratio of the WL delays without (D_{org}) and with (D) the reconfiguring inverter is plotted in Fig. 5.

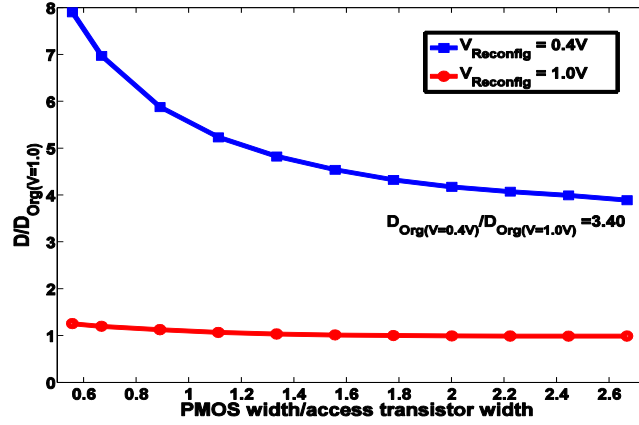


Fig. 5: Effect of PMOS size of the ReconfigInv on WL delay. D is WL delay through the reconfiguration network, and D_{org} is the original WL delay.

For a given V_{low} , the delay of the reconfigurable WL for different PMOS width is estimated, and the width that provides the minimum WL delay is estimated for the low-voltage mode. The maximum allowable PMOS width can be obtained from the cell layout. IBM 130nm CMOS technology is considered to estimate the area overhead. The layout of one row of the MUX group shows that the inverter area can be kept under “half-cell” width resulting in ~6% overhead for the core SRAM array as shown in Fig. 6. The smaller one between the maximum PMOS width from the layout and the optimal PMOS width for the minimum delay is used as the final PMOS width. A minimum size NMOS is used to reduce energy. The ReconfigInvs for pre-charge and column-selection network can be sized using the same principle.

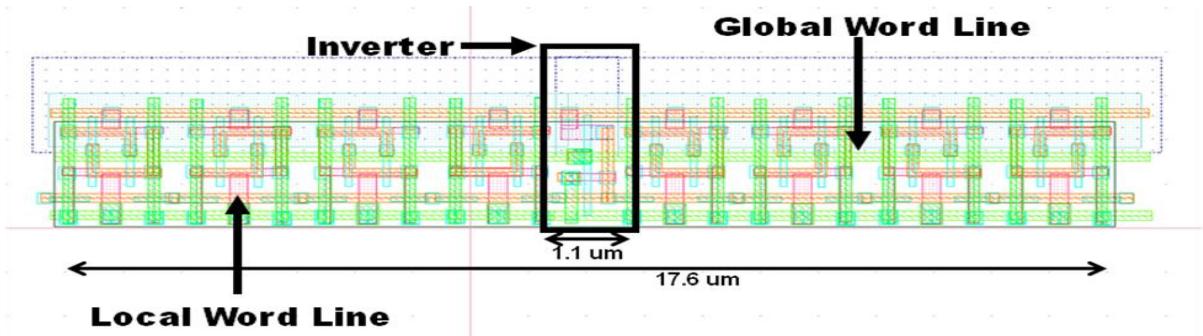


Fig. 6: Layout of one row of the proposed reconfigurable SRAM showing the reconfiguring inverters.

Voltage-Reconfiguration Network

The reconfiguration of the array is performed by changing the voltage levels for a number of LOBs to high or low voltage. The voltage-reconfiguration network consists of two PMOSes – one connected to the low voltage and the other to the high voltage. The gate voltage of the PMOSes for the j -th bit is controlled by the j -th select signal (sel_j). To reconfigure the j -th bit to the low voltage, sel_j is set to “0”. Note that this architecture also allows uniform voltage scaling, which can be performed by setting $sel_j=0$ for all the bits and changing the V_{low} . The switching PMOSes in the voltage-reconfiguration network for the array and pre-charge devices need to ensure a low voltage drop. A high drop will reduce the effective cell supply and can degrade the read margin. However, as the same PMOS also acts as the supply for the ReconfigInv, the wordline and cell-supply voltage remain the same, which reduces this effect. The width of the PMOS marginally impacts the switching speed of the ReconfigInvs, which has minimal effect on access and write failures.

Reconfiguration Length

Implementation of a given reconfiguration length can be performed by putting multiple bits in the same voltage network. For example, $RL=2$ implies that MUX Groups for Bit 0 and Bit 1 will share the same cell supply and bitline-voltage network, the same local wordline, and the same reconfiguration inverter. A short reconfiguration length improves power saving and provides more room for accuracy-power tradeoff. Note that RL does not have a direct impact on the design overhead for the array, but increasing RL proportionately increases the size of ReconfigInvs. RL also does not impact the performance or failure rates of the cell. The reconfiguration length determines the number of reconfiguration bits. Reconfiguration bits are the input signals that are decoded to generate control signals for the voltage-configuration

network. The number of reconfiguration bits is a key limiting factor for small RL length for 16 or 24-bit image storage. For example, if a 16-bit image and $RL=1$ are considered, a separate control signal is required for voltage-configuration network for each bit, which can be generated by decoding 4 reconfiguration bits. If $RL=4$, 4 different control signals are needed. 1st signal will be used to control voltage-configuration network for Bit 0 to Bit 3, 2nd signal for Bit 4 to Bit 7, 3rd signal for Bit 8 to Bit 11, and 4th signal for Bit 12 to Bit 15.

Reconfiguration Time

The reconfiguration of the array essentially implies dynamic changes in the voltage of MUX Groups from low to nominal level or vice-versa for a number of bits. The reconfiguration is performed in runtime and can always be revoked. The array can move between a high-error and a low-error mode in runtime. The error characteristics in the low-error mode (i.e., number of low-voltage bits) can also be changed in runtime. The choice of the size of PMOS devices in the voltage-reconfiguration network determines the reconfiguration time (i.e., the time required to change the voltage level of a bit). This is because the supply line for the MUX group associated with a bit needs to be discharged from high to low voltage or charged from low to high voltage during reconfiguration. Ideally, the voltage-reconfiguration network for a MUX Group can be designed using two properly sized PMOS devices. However, to improve the voltage stability, a distributed network is considered (i.e., one voltage-reconfiguration network with two PMOS devices per column). The time required to change the common voltage (i.e., the cell supply, the supply of the ReconfigInvs, the supply of pre-charge transistors, and the write driver) is estimated considering the distributed PMOS network per column. The capacitance of the supply line is estimated considering 256 cells, pre-charge PMOSes, write drives, and 256 ReconfigInvs. Note that ReconfigInvs are shared across a MUX Group to provide a worst-case estimate. The metal

capacitance of the shared supply line is estimated considering height of the cells and $0.2\text{fF}/\mu\text{m}$ metal capacitance and included in the analysis. The PMOSes in the voltage-reconfiguration network per column have 5X width of the cell PMOSes, which results in an area overhead of approximately one SRAM cell per column of 256 cells ($\sim 0.4\%$ area overhead). With this additional area and considering predictive 70nm technology, the estimated reconfiguration time is $\sim 4\text{ns}$. For a 250MHz clock cycle, 4ns corresponds to one clock cycle of reconfiguration time. Since all the columns are reconfigured in parallel, the total reconfiguration time for the array is also the same. If a 256×256 pixel image is read and reconfiguration is performed for every image, the performance overhead is negligible. The access protocol will require an additional clock cycle per image for reconfiguration. The reconfiguration time can be reduced at the cost of additional area.

3.3.3 Array-Power Estimation

The power dissipation in the memory array depends on the read or write operations. In this section, we present the power model for the two operations and explain the origin of power saving in the proposed method.

Read-Access Power

During a read operation, energy is dissipated because of switching of the wordline and bitlines.

The wordline power can be computed as

$$\begin{aligned}
 E_{WL} &= E_{driver} + E_{ReconfigInvs} \\
 E_{driver} &= (N_{COL} C_{bitWL} + N_{bit} C_{inv}) V_{high}^2 \\
 E_{ReconfigInvs} &= \left[N_{mux} (C_{bitWL} + 2C_{ax}) \right] \\
 &\quad \times \left[(N_{bit} - L_{bit}) V_{high}^2 + L_{bit} V_{low}^2 \right].
 \end{aligned} \tag{7}$$

For all bits operating at high voltage, this energy is marginally higher than the wordline driver

energy for regular array. Eq. (7) can be used to size the ReconfigInvs to minimize energy dissipation, while the delay is kept under an acceptable limit. The energy dissipated in the pre-charge and column-selection network can also be estimated similarly. The bitline energy depends on voltage swing at the bitline, which in turn depends on the cell and wordline voltage. The bitline power can be estimated as

$$E_{bitline} = N_{row} C_{bitBL} \left[(N_{bit} - L_{bit}) \Delta_{bit_high} V_{high} + L_{bit} \Delta_{bit_low} V_{low} \right] \quad (8)$$

where C_{bitBL} is the bitline capacitance per cell [includes interconnect and device (junction + overlap) capacitance of access transistor], and Δ_{bit_high} and Δ_{bit_low} are the bit-differential developed during reading for the nominal and low voltage respectively. For a medium operating frequency, the nominal supply voltage results in a large Δ_{bit} resulting in low access failure. However, as explained in Section 3.2, it does not provide an opportunity for supply reduction for the entire array as the cell failure is dominated by disturb failures. Eq. (7) and (8) clearly suggest that operating a larger number of bits at a low voltage reduces more power.

Write Power

During a write operation, energy is dissipated in wordline drivers, write driver and pre-charge devices. The write energy dissipated at the bitline can be estimated as

$$E_{bl_write} = C_{BL} \left[(N_{bit} - L_{bit}) V_{high}^2 + L_{bit} V_{low}^2 \right] \quad (9)$$

The above models can provide an initial estimate of the power saving considering different voltage levels and number of bits in the low-voltage mode. The detailed power calculation requires measuring different components from circuit simulations.

Leakage Energy

Reducing voltage at LOBs also reduces the active leakage in the array (i.e., leakage energy dissipates in the unselected cells in the accessed array). The leakage power can be estimated as

$$P_{leakage} = N_{mux} (N_{row} - 1) \left[V_{high} (N_{bit} - L_{bit}) I_{cell} (@ V_{high}) + V_{low} L_{bit} I_{cell} (@ V_{low}) \right], \quad (10)$$

where I_{cell} is the cell-leakage current at different voltage levels, which is computed using a circuit simulation. Since an active array can be at a high temperature, the leakage energy is expected to be dominated by the subthreshold current. This analysis does not consider the sense-amplifier power and the decoder logic power. Decoder power can be reduced by operating it at a low supply voltage, but this operation requires a level-converting-wordline driver. The sense-amplifier power can be reduced only if the data bus is allowed to be driven at a low voltage.

3.4 System-Level Image-Quality Estimation

The efficient and accurate analysis of the effect of spatial voltage scaling on application quality is a key concern for the feasibility. As explained in Section 3.2, the parametric failures in memory can occur because of different physical mechanisms. The different failure mechanisms have varying dependence on circuit parameters. Moreover, the effect of cell failures on image-quality degradation depends not only on the overall cell-failure probability (i.e., number of faulty cells in an array) but also the fault locations within an array. The nature and extent of process variations within the cell that can cause different types of failures are also different. For example, while write failures are caused by strong PMOS and weak access transistors, disturb failures are caused by strong access, weak pull-down, and weak PMOS transistors. In other words, the fault locations for different types of failures can be different. Hence, the effect of memory-induced error on the image quality will depend on the type of memory access (i.e., read or write) and the

location of corresponding faults (i.e., disturb or access failures for read access and write for write access). The analysis of the effectiveness of spatial voltage scaling on SRAM needs to consider circuit-level details, number and location of the faults in an array, and the types of access (i.e., read or write) to the memory array. Such an analysis cannot be accurately performed using a conventional yield-estimation framework [6-8] or using a lumped bit-error-rate number [13]. A full-chip Monte-Carlo simulation is necessary to assign random V_{th} values to all transistors in the array, and perform the read or the write operations. However, such a method significantly increases the simulation time and cannot be used for exploration of system-level design parameters, such as the number of bits in the low-voltage domain and the values of V_{high} and V_{low} . The system-level image-quality estimation requires a unique fault-simulation framework that connects circuit-level estimation of failure rates to system-level applications, which determines the types of memory accesses. The last part is more important as the error introduced by the memory faults in an image can propagate over time. A system-level fault simulator has been developed to address the above mentioned challenge as shown in Fig. 7.

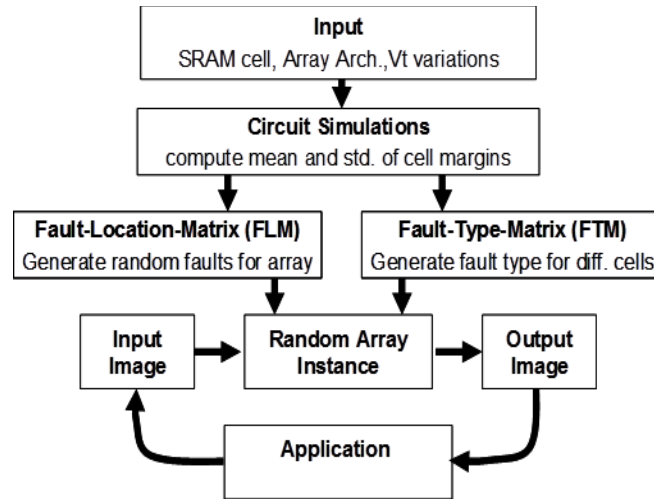


Fig. 7: Simulation framework.

- **Step – 1: Failure probability simulations:** The circuit simulations are performed to

estimate the read margin, Δ_{BIT} , T_{write} , and offset distributions at different voltages. The distributions are fitted to a normal density, and mean and standard deviations are estimated ($1/T_{\text{write}}$ was considered for the write time). Note that the failure probability simulation represents the characteristics of technology, cell design, and array design. For the proposed array, simulations need to consider the ReconfigInvs and voltage-switching network, which can modify the cell margins.

- **Step – 2: Creation of fault location map (FLM):**

- Given an array, random values for RM, Δ_{BIT} , and T_{write} values are generated for each cell. Random offset values are generated for sense-amplifiers. One generation of the random values essentially represents one chip. For different bits operating at different voltages, the random values from the distribution corresponding to that voltage are generated.
- Depending on the generated values, the fault location map (FLM) for the array is generated considering the failure models as discussed in Section 3.2. For a multi-voltage domain, fault maps for different domains are prepared and merged together to create the overall FLM for the array. A separate FLM is created for each operation. FLM represents which cells in an array are faulty in a particular instant of SRAM die generated during one Monte-Carlo simulation run.

- **Step – 3: Creation of fault-type map (FTM):** Since the failures originate from device mismatches [6, 7], for a given failure mechanism, cells failing while storing “1” is less likely to fail while storing “0”. To reflect this fact, the framework randomly generates a weak “1” or weak “0” indicator for each cell. Note that a weak “1” or weak “0” does not represent the failure, but represents that the cell is more likely to fail for bit “1” or bit “0”. A fault-type map of an array represents which cells in an array are more likely to fail while storing “0” and which cells are

more likely to fail while storing “1”.

- **Step – 4: Creation of an SRAM array with faults:** A particular array (i.e., a particular chip) is characterized by three FLMs and three FTMs, and each array corresponds to one failure mechanism (i.e., access, disturb, or write failures). Given high and low-voltage levels and the number of bits in the low-voltage domain, a random array instance can be created using Step-2 to Step-4. The created array instance contains technology and circuit, and μ -architecture information.
- **Step – 5: System-level fault simulations:** Consider a memory-access operation such as the image read or the image write. For a given operation, each bit in the image is mapped to one memory cell in the 2D array. Different bits of a single pixel are mapped to different MUX groups (i.e., column multiplexing). While mapping the image, the random faults in array modify the image (i.e., introduce error). If the mapped location of the bit indicates a fault in the FLM and the bit value matches with the value in corresponding location in the FTM, the bit is inverted to indicate an error. The logical operation is defined as

$$\begin{aligned}
 bit'[i,j] &= \overline{FLM[l,j]} \cdot bit[i,j] + FLM[i,j] (\overline{bit[i,j]} \cdot \overline{FTM[l,j]} + \\
 &\quad \overline{bit[l,j]} \cdot \overline{FTM[l,j]}) \\
 &= \overline{FLM[l,j]} \cdot bit[i,j] + FLM[i,j] \cdot \overline{FTM[l,j]},
 \end{aligned} \tag{11}$$

where i, j , represent the bit location in the digitized image matrix and corresponding locations in the FLM and FTM matrix, $bit[i, j]$ and $bit'[i, j]$ represent the original and modified bit values, respectively. The logical nature of the operation can be used to perform a bit-wise operation on the integer image matrix by converting the FLM and FTM to integer matrices (considering column multiplexing), $FLM^{(int)}$ and $FTM^{(int)}$. Therefore, the fast bitwise logical operators can be used to significantly reduce the simulation time. This framework is particularly helpful for an

application-level simulation, which uses the system-level simulator to calculate the quality degradation. The effect of a memory-failure mechanism on an image (A) is therefore computed as

$$A_{op} = or \left[and \left(A, not \left(FLM_{op}^{(int)} \right) \right), \right. \\ \left. and \left(FLM_{op}^{(int)}, not \left(FTM_{op}^{(int)} \right) \right) \right], \quad (12)$$

where different logical operations indicate the element-by-element bitwise operation between two matrices. To estimate the degradation in the quality of an image caused by random faults, an unit operation is defined as a sequence of three operations, namely, image write, followed by a first image read (which creates the disturb failures), and followed by a second image read (which captures access failures in the disturbed image) as

$$\begin{aligned} A_{write} &= write(A_{in}, FLM_{write}^{(int)}, FTM_{write}^{(int)}) \\ A_{disturb} &= read(A_{write}, FLM_{disturb}^{(int)}, FTM_{disturb}^{(int)}) \\ A_{final} &= read(A_{disturb}, FLM_{access}^{(int)}, FTM_{access}^{(int)}) \\ MSSIM(A_{in}) &= compute_ssim(A_{in}, A_{final}) \end{aligned} \quad (13)$$

where A_{in} is the input image, A_{write} , $A_{disturb}$, and A_{final} is the images after write, first read, and second read operation.

- **Worst-case simulation:** We also create a worst-case FLM, which ignores the failure mechanism of the cells and considers a cell faulty if it fails because of any mechanism (i.e., bitwise ‘or’ operation of three elemental FLMs). This approach also ignores the fault-type information. The worst-case FLM can be used for design-space exploration as the FLM captures the worst-case scenario. The system-level fault simulator is used to evaluate the effect of the proposed architecture on an application.

3.5 Simulation Results

In this section, the simulation results are presented to explain the effectiveness of the proposed scheme in saving power, and the achievable accuracy-energy tradeoff is discussed.

3.5.1 Circuit Simulations

The simulations in predictive 70nm technology are performed to evaluate the key circuit parameters for the proposed scheme. First, the effects of supply voltage on access, write, and disturb failure probability of a given SRAM cell ($W_{PUP}:W_{AX}:W_{PD} = 1:1.125:1.625$, where W_{PUP} is the width of the PMOS pull-up device, W_{AX} is the width of the access device, and W_{PD} is the pull-down device) are evaluated considering the ReconfigInvs and the PMOS-switching network. The effect of increased WL delay at the low voltage was observed for access and write failures, but the effect of PMOS-switching network was negligible. Read margin, access margin, and $1/T_{write}$ was observed to reasonably follow a normal density down to 0.4V. The mean and standard deviations of read margin, access margin, and $1/T_{write}$ were estimated from the circuit simulations at different supply voltages and used in the system-level fault simulator. The circuit-level simulations were also performed to estimate the power dissipation of the key circuit components at different voltages following the discussion in Section 3.5.3. Finally, the component powers are used to estimate the overall array power.

3.5.2 System-Level Simulation of Images

The effect of the proposed architecture is evaluated on a standard 8-bit grayscale test image suite available in [17] considering 250MHz of operation. The power saving was computed with reference to a regular array [all bits at the nominal voltage (1V)] to show the effect of spatial voltage scaling and voltage reduction for all bits. The average of the read and write power is considered as the array power in the analysis.

Effect of Voltage Scaling on Image Quality and Power

Fig. 8 shows the degradation of image quality (computed using the MSSIM) and corresponding power saving for different images for different low-voltage levels after performing the unit operation described in Eq. (13). Significant power saving can be obtained with a graceful degradation of image quality. As expected, $L_{\text{bit}}=8$ provides more power saving at the cost of quality. Using $L_{\text{bit}}=4$, ~45% power saving can be obtained compared to a regular array with 10% reduction in quality. Considering the fact the memory power in multimedia applications can be as much as 50%, the proposed architecture with $L_{\text{bit}}=4$ and $V_{\text{low}}=0.4\text{V}$ can result in overall 23% saving in the system power. Using $L_{\text{bit}}=6$, the memory power saving increases to ~75% resulting in ~37% saving in the system power. Power saving is 20% higher than the saving achievable by reducing voltage of all bits (“blind scaling”) at the same degradation level.

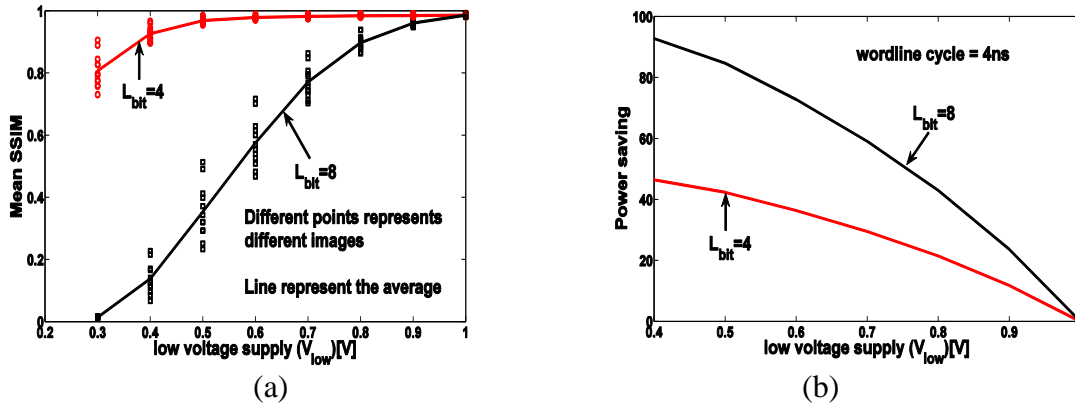


Fig. 8: Effect of supply voltage scaling on (a) image quality and (b) power saving.

The reproduced images clearly demonstrate the advantages of using the proposed accuracy-aware SRAM with spatial voltage scaling over blind voltage scaling as shown in Fig. 9.



$$V_{dd} = 0.4V$$

(a)



$$V_{high} = 1.0V, V_{low} = 0.4V$$

(b)

Fig. 9: Image with Vdd scaling (a) all bit at 0.4V and (b) $L_{bit} = 4$ and $V_{low} = 0.4V$.

Reconfiguration: Accuracy-Energy Tradeoff

The effect of reconfiguration of L_{bit} at a given low-voltage level is considered. Fig. 10(a) shows the MSSIM and power saving for different number of bits in the low-voltage mode. The results considering two different V_{low} values are shown. Increasing the number of reconfiguration bit degrades the quality with an increase in power saving over the regular array. The quality degrades slowly till the 6th bit, but reconfiguring the 7th and 8th bit to the low-voltage mode can result in significant error. Further, a high V_{low} provides a high room for reconfiguration with a low power saving in each configuration. Fig. 10(a) shows that the proposed array provides a unique way to perform the run-time tradeoff with only two voltage levels. Compared to previous work [13], reconfiguration provides more efficient power saving since the number of low-voltage-bits (L_{bit}) can be increased in runtime depending on error requirement or image characteristics. Note that the proposed approach does not aim to improve the design yield as in the case of redundancy or correct errors in runtime as in the case of error correction coding (ECC).

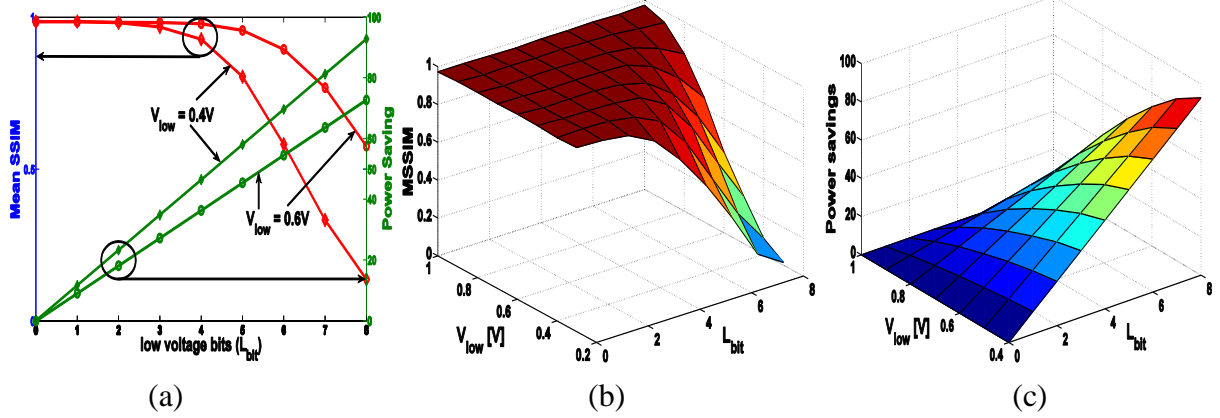


Fig. 10: Reconfiguration: (a) image quality and power tradeoff for a given V_{low} , 3D plots showing the co-reconfiguration of L_{bit} and V_{low} and its effect on (b) image quality and (c) power saving.

Design of V_{low}

The results from previous sections suggest that a co-optimization of the number of bits and voltage level can provide a better insight into the problem. These simulations were performed considering the worst-case FLM for a given image. If both L_{bit} and V_{low} can be reconfigured, more power saving can be obtained as shown in Fig. 10(b) and 10(c). However, a more practical approach is to select a low-voltage level and use L_{bit} for reconfiguration as fine-grain change in the supply voltage can significantly increase the design complexity.

Effect of Chip-to-Chip Variation in Fault Locations

Since, for a given voltage and L_{bit} value, different chips are going to have different fault locations, a Monte-Carlo simulation is performed using the system-level simulator. Each instance of the Monte-Carlo simulation results in a different FLM and FTM, even for same value of V_{low} and L_{bit} . The image-quality degradations for different images in test suite are evaluated. A consistent improvement in image quality was observed and compared to the case with all bits going to low voltage. Fig. 11 shows that with uniform voltage scaling one can have large variation in the image quality degradation from one fabricated chip to another, even though the

extent of the local process variation remains same. This is due to the randomness in the location of the faults.

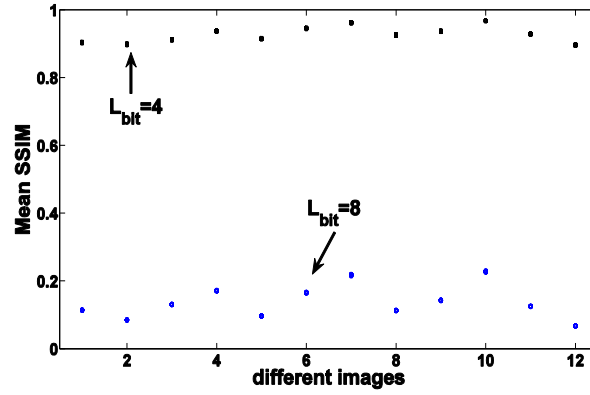


Fig. 11: Quality degradation of different images considering multiple MC runs.

Effect of Transient Noise

In the previous analysis, the manufacturing variations are considered. However, cells in the weak corner (may not be failing), can fail because of transient noise such as the thermal or supply noise at different times. Such noise may not scale with the voltage. To capture transient noise, an image was read repeatedly, and a Normal variation is applied to the RM of the cells (over the RM obtained after manufacturing variations) during each read operation, which results in transient disturb failures to the image during each read. The MSSIM after each read operation is estimated. As “blind scaling” makes HOB cells “weak”, an image becomes more susceptible to transient noise as shown in Fig. 12. Therefore, even if the initial image quality is similar for ($L_{bit}=4$, $V_{low}=0.4$) and ($L_{bit}=4$, $V_{low}=V_{high}=0.8V$), the quality is degraded at a fast rate for “blind scaling”.

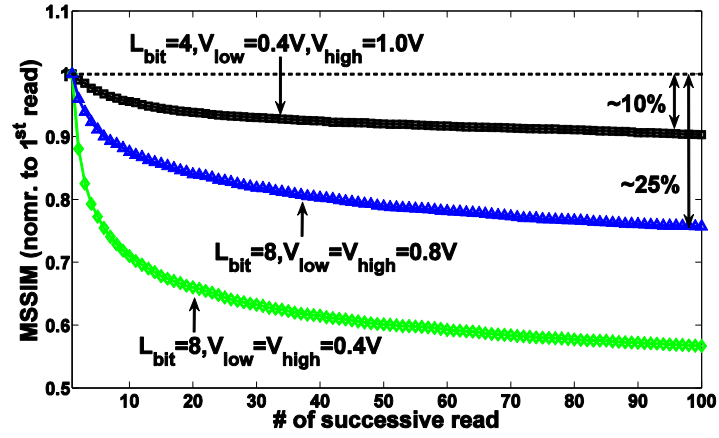


Fig. 12: Impact of transient noise on repeated reading of same image (MSSIM normalized to that for 1st read).

CHAPTER IV

DESIGN METHOD TO CHARACTERIZE AND COMPENSATE FOR PROCESS VARIATION IN 3D ICS

4.1 Need for Testing and Characterization of Process Variation 3D ICs

The through silicon vias (TSVs) in a 3D stack are the channels for transferring signals between different tiers in a 3D stack. The functionality of a 3D IC strongly depends on the fidelity of signals through TSVs [18-22]. As the TSV process is not a perfect one, defects can be created while forming the TSVs before bonding (assuming a via-first process) or while bonding different dies together [18-29]. The defect can be created by a short through the oxide surrounding the TSVs, which results in finite resistance between the TSV and the substrate. The open defects or ruptures can also be created during TSV growth. The non-conformal growth of the insulator also creates defects or variation in TSV properties. At the post-bond stage, the defects can be created because of the variation in the resistance of the bonding material or TSV misalignment.

The electrical effects of such defects are partial or complete degradation of signal fidelity through TSVs. A short through the oxide creates a resistive path through the oxide as shown in Fig. 13. Assuming the substrate surrounding the TSVs is connected to ground, the short is a low-resistive path between the TSV and ground. Likewise the open defect and the variation in bonding resistance, TSV misalignment impacts the resistance through a TSV. When the TSV is driven by a driver, the signal swing and slew at the receiver end can vary significantly because of short defects, which results in either complete or partial signal degradation. The complete degradation is caused by a strong short, while a weak short results in partial degradation.

TSV faults need to be detected before (pre-bond test) and after bonding (post-bond test) [18-

29]. The need for post-bond detection is obvious to ensure that faulty 3D ICs are not shipped to the customer. On the other hand, the pre-bond detection can help to screen the dies with defects before bonding and can help reduce potential yield loss by bonding a faulty die with a good one [21-23]. However, only detection of the faults may lead to high pessimism in design yield as all short or open defects may not be critical for signal fidelity. Therefore, the characterization of electrical strength of the individual defects is important.

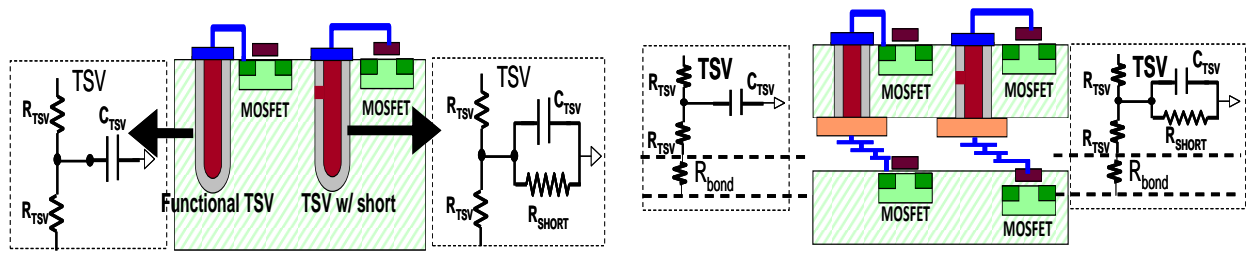


Fig. 13: Schematic illustration of TSV shorts at the pre-bond stage (left) and variation in the resistance at the post-bond stage (right).

4.2 Challenges of Design for Testing and Characterization Structure of the TSVs

Testing and characterization of the TSVs have significant challenges, particularly before bonding [20-22]. The TSVs are too small for test probes, and one cannot afford to include a large number of probe pads for testing. Hence, one needs to design built-in test structures, which can characterize the TSVs before bonding. Pre-bond test structures need to satisfy additional requirements. First, the test structures used for pre-bond testing should be designed such that it can also characterize TSV defects after bonding. Second, even if the TSV has a defect, unless it is a catastrophic one, the signal can propagate through the TSV with the degradation in amplitude and slew. If the degraded signal can be recovered at the receiver end of the TSV, it is possible to repair the TSVs with moderate defects while maintaining the required system-level signal fidelity and improving the overall design yield. Therefore, the test structures should also

be able to function as signal-recovery circuits that can repair the TSVs with moderate electrical degradation. Third, for accurate characterization and recovery, the test structures will be required for individual TSVs. As the number of TSVs in a 3D IC increases, it is critical that the structures should be simple and low power. Distinguishing between weak and strong defects requires testing for the analog properties of TSVs and correlating the properties to signal quality. For small number of TSVs, direct analog measurement is possible, but as the number of TSVs is large, it is important that test structures should be able to create digital signatures of the analog nature of the defects. The digital signatures need to be stored on a chip and later read out for test, characterization, and recovery. Finally, it is important to analyze the power, performance, and area overhead associated with the built-in test structures considering full-chip analysis of 3D ICs. Such analysis requires incorporating the test structures in the physical design of a 3D IC and performing the detailed power and performance analysis.

4.3 Electrical Effects of TSV Defects

This section presents the electrical impact of the TSV defects on the signal quality. The electrical effect is addressed considering the resistive shorts through the surrounding effect. Fig. 14 shows a typical scenario, in which a TSV with short defect is driven by an inverter in one die, and the signal is received by another inverter on the second die. The signal degradation at the receiver end is of a primary concern for correct functionality of the 3D ICs. Consider variation in the resistance of the short because of variation in the diameter of the short. The voltage at the receiver end can vary significantly, which results in either complete or partial signal degradation. Based on the extent of the signal degradation, the TSVs are partitioned in three categories. If a low-resistance short exists, V_{TSV} will be very low. This TSV is referred to as un-repairable or *bad* TSVs. The TSVs, for which V_{TSV} is very high ($>90\% V_{DD}$), are considered as *good* or

defect-free TSVs. We also define a third category of TSVs, referred to as *repairable* TSVs, which corresponds to shorts with moderately high resistance such that signal degradation is within an acceptable limit (e.g., 50% of V_{DD}). During signal propagation through such TSVs, the voltage at the receiver end will make a logical transition between “0” and “1”, but experience a reduced swing. A low voltage swing leads to high noise susceptibility, short-circuit power, and delay at the receiving gate. Fig. 14 shows the delay, average power, and signal swing at the receiver end of a TSV with different short resistances. For very high resistance of the shorts (i.e., good TSVs), the signal swing is close to ideal. As the resistance reduces, the signal swing reduces gradually resulting in corresponding increase in signal delay through the TSV and average power of the driver/receiver combination. The TSVs with such intermediate resistances of the short belong to the *repairable* TSVs category. We refer to the TSVs as repairable because if the signal swing can be recovered at the receiver end, the TSV will be functioning possibly with a higher delay than the *good* TSVs. If the resistance is very low, the signal fails to make a transition. After the bonding, variation in TSV resistance causes the degradation of signal swing through TSV channel.

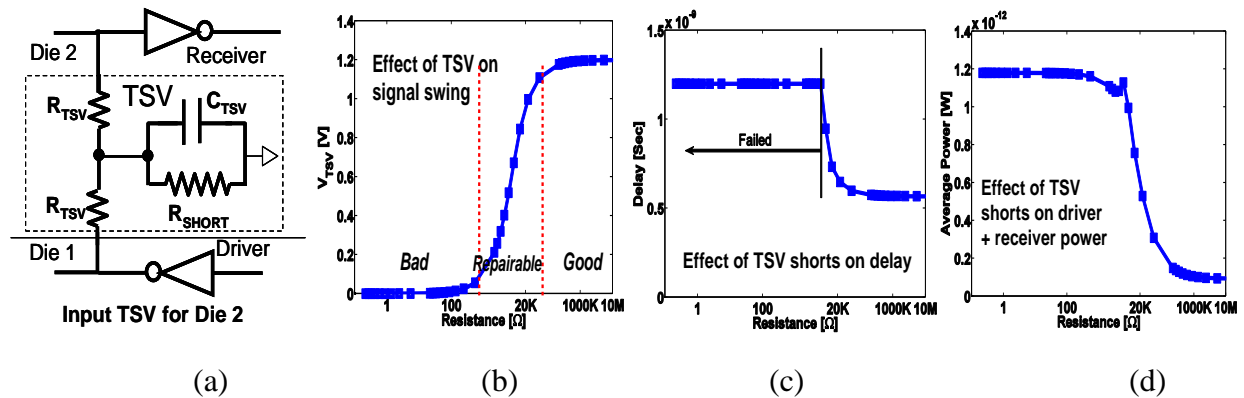


Fig. 14: Effect of TSV short: (a) driver-receiver combination, (b) signal swing, (c) delay, and (d) driver and receiver power.

Fig. 15(a) shows a scenario where resistance variation can occur between a TSV and a driver/receiver because of a weak open, misalignment, and bonding resistance. The combined effect is collectively modeled as variation in net TSV resistance. For small variation in this TSV resistance, signal swing is close to ideal. As the variation increases, signal swing through TSV starts to reduce, and eventually signal fails to function properly. Variation in TSV resistance degrades signal swing as shown in Fig. 15(b), impacts signal slew as shown in Fig. 15(c), and increases the delay of signal through TSV as shown in Fig. 15(d).

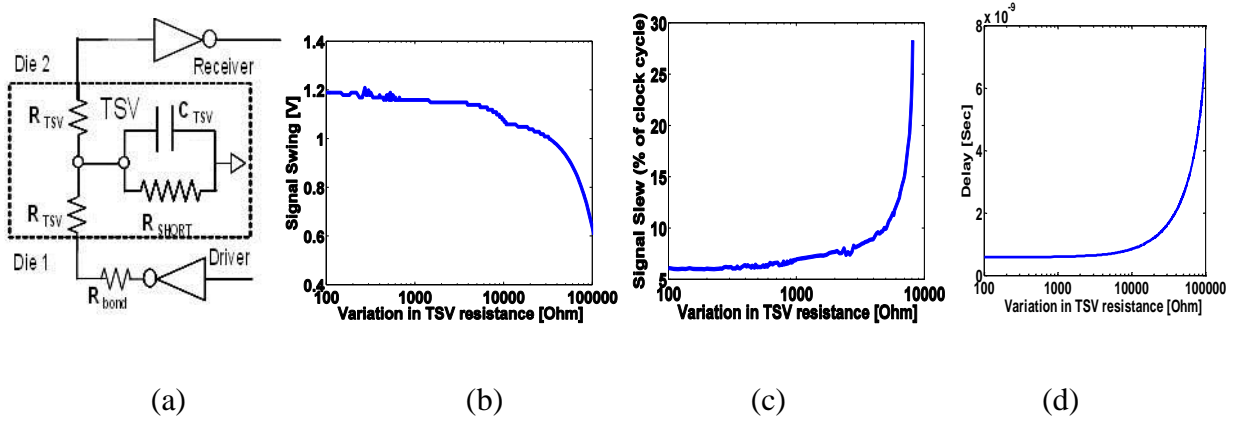


Fig. 15: Post-bond TSV resistance: (a) driver-receiver, (b) effect on signal swing, (c) signal slew, and (d) signal delay.

4.4 Test and Signal-Recovery Structure for TSV

This section discusses the proposed structure and the operating principle. After the bonding, the driver and the receiver across a TSV will be at two different tiers. Hence, during pre-bond test, it is important to design a test circuit that can mimic the voltage degradation through the TSVs. However, a test structure needs to mimic this degradation by using all devices in the same tier. The objective of the proposed test structure is to first place individual TSVs in one of three categories during pre-bond test: *bad*, *repairable*, or *good*. If a *bad* TSV exists (assuming non-redundant TSVs), the die is detected as a faulty one, which is not used in bonding and adds to

yield degradation. However, for the repairable TSVs, the test structure reconfigures itself as a signal-recovery circuit for post-bond normal operation. This structure allows reducing the overall yield degradation at the expense of the marginal delay and power overhead. For the *good* TSVs, the test structure allows direct signal transfer from the driver to the receiver without signal recovery. After bonding, the test structure retests individual TSVs to capture the effect of variations in resistance of TSVs. Note that the test structure for repairable TSVs will continue to operate in the recovery mode even after bonding. If the variation in resistance of the TSVs is very high and cause signal degradation even through TSVs detected as defect-free at the pre-bond test, the test structure for such TSVs also reconfigures itself to the signal-recovery circuit.

4.4.1 The Basic Structure and Operating Principle

Pre-bond Test Mode for Input TSV

A TSV test inverter (TTI) is connected to each TSV as shown in Fig. 16. During the test, the input of the TTI is held low. This input forms a resistor-divider structure between the PMOS resistance of the TTI and the resistance of the TSV short. The resistances of the short (R_{short}) and the TTI determine the voltage at the TSV-TTI junction (V_{TSV}). Depending on the value of the R_{short} , the V_{TSV} will vary. V_{TSV} is next compared against a reference voltage and sampled into a scan flip-flop connected to the comparator. The reference voltage is selected such that it represents an “acceptable” signal quality (i.e., $>50\%$ of V_{DD}). The TSVs with low-resistive shorts will have values less than the reference voltage, while the defect-free TSVs will have very high voltage. The scan flip-flops (FFs) of the TSVs form a scan-chain. The output of the comparator connected to a TSV captures the extent of TSV short in a digital form. At the end of the test, the values stored in the scan flip-flops are scanned out to locate the faulty TSVs. If such TSVs exist, the die is categorized as a faulty one. The above test is performed again with a high reference

voltage ($\sim 90\%$ of V_{DD}). The scan FFs, which indicate faults with this high reference voltage, correspond to *repairable* TSVs.

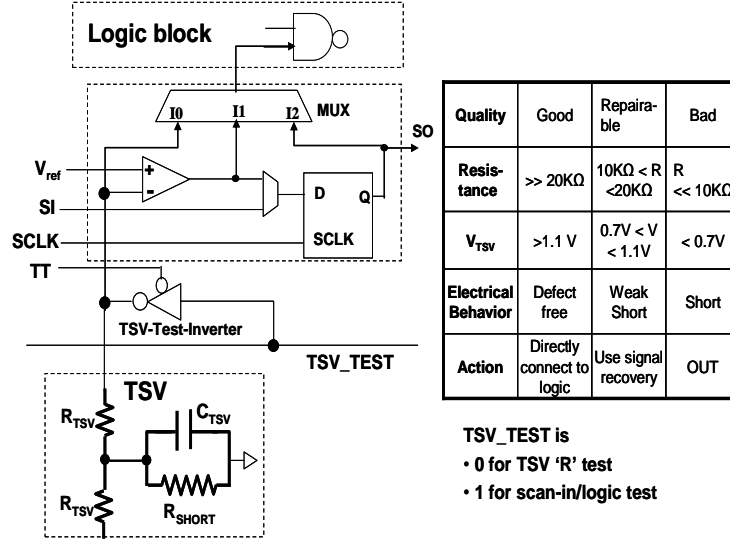


Fig. 16: The basic test structure and characterization policy.

Signal Recovery during Normal Operation (Input TSV)

For the repairable TSVs, the test circuit reconfigures itself to connect the output of the comparator to the input of the logic gates (instead of directly using the TSVs) during normal operation. The comparator is designed such that during normal operation it functions as level-converter circuits and recovers the signal degradation. As for the input TSVs, the test circuit resides between the TSVs and the input logic gate.

Pre-bond Testing and Signal Recovery for the Output TSV

The test structure for pre-bond testing of the output TSV is similar to that of an input TSV and uses the same test and recovery circuit. The TTI is connected to the TSV, and the V_{TSV} is sampled by the test circuit. The basic operating principle is also similar. The only difference is

that, instead of an output MUX, the input MUX, scan in (SI), and the logic output need to be extended. The primary difference in this case is that the signal recovery needs to be performed in the second die as shown in Fig. 17(a). This difference will require re-running the TSV test after bonding where TTI will drive the bonded TSV. The V_{TSV} developed at the receiver end in the second die will be compared against the threshold to activate or deactivate the signal recovery.

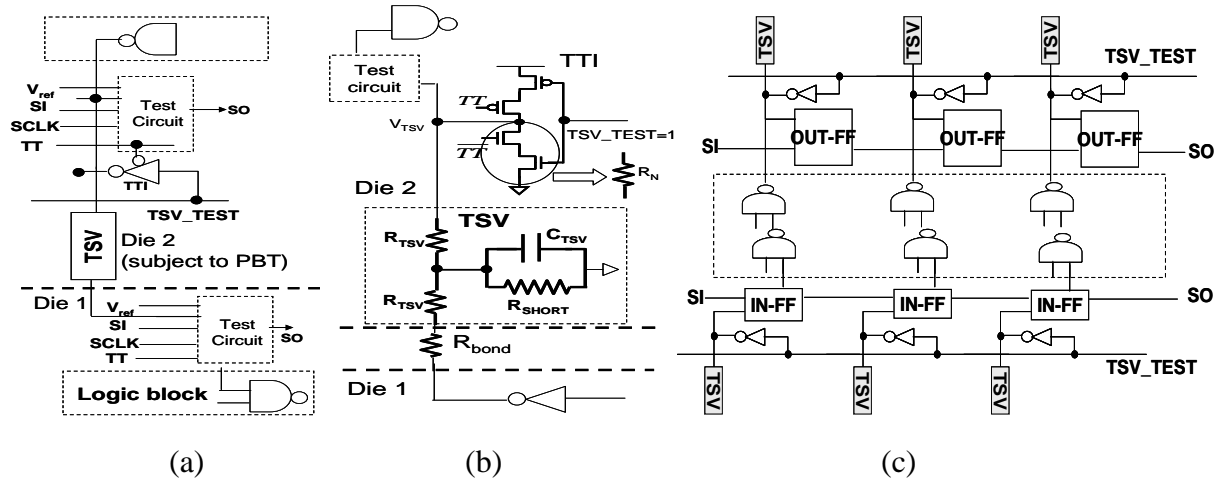


Fig. 17: Test structure for (a) pre-bond test of output TSV, (b) post-bond characterization, and (c) the logic diagram of the overall test structure.

Post-bond Testing and Signal Recovery

Along with the short defects in the TSVs, the effective resistance of the TSV can also vary because of weak open, resistance of the bonding material, or misalignment in a 3D process. The proposed test structure is also used to characterize variations in the resistance of TSV. If required, the test structure can perform signal recovery as shown in Fig. 17(b). If a TSV is detected as a repairable one in the pre-bond stage, it will automatically be configured in the signal-recovery mode. Further, as such TSVs will have the moderate short, the voltage developed at the TSV end will vary with a variation in the resistance of TSVs. Hence, our test approach is to activate the TSVs from the driver layers and sample the TSV voltage at the receiver layer using the proposed

test circuit. However, for good TSVs with very high resistive short, a variation in the resistance of TSVs will not be reflected in the DC voltage change of the TSVs. Since the TSV is connected to the gate of a device in the receiver end, no current path to ground exists. To enable the characterization of variation in the resistance of TSVs in such scenarios, we propose to assign TSV_TEST=1 for the TTI in the receiver tier. The configuration activates the NMOS device in the TTI in the receiver end and creates a current path from the driver in Die 1 to the ground. The NMOS will be designed with a long channel (or multiple NMOS in series) to ensure that the voltage drop across the NMOS is very high if variation in the resistance of TSVs is very low. Otherwise, depending on the variation in the resistance, V_{TSV} will vary. The developed V_{TSV} into the test circuit is sampled to detect whether the variation in the resistance of TSVs is higher than a given limit. A high variation in resistance will degrade the signal slew (for defect-free TSVs) and signal swing (for TSVs with short of moderate resistance). If a low V_{TSV} is developed during this test, the signal-recovery circuit is activated to improve the signal swing and slew.

Integration with Scan Architecture for Pre-Bond Functional Test

This chapter primarily focuses on the TSV test. However, the scan flip-flops used in the TSV test can also perform as a scan chain for pre-bond logic test as shown in Fig. 17(c). This overall structure enables to perform pre-bond functional tests on the partial circuits in each die. The required test vectors are scanned in to assign desired signal values to the logic inputs, which are connected to the TSVs. Similarly, the logic outputs connected to the TSVs are also sampled at the output TSV scan chain and analyzed to detect whether logic defects exist in the partial die. It is imperative that the existing flip-flops in each die will also be converted to scan FF to enable full or partial scans.

4.4.2 Detailed Circuit Design of the Test and Recovery Circuit

Fig. 18 shows the detailed circuit schematic of the proposed test structure, and Table. 1 shows the test control and methodology. The operation is explained considering the pre-bond test condition. Note that the scan flip-flops used in the test structure do not function as a flip-flop during the regular operation. The proposed flip-flop allows innovating in the design of the proposed structure. The heart of the proposed structure is a differential sense-amplifier-based flip-flop. The flip-flop is designed only with the PMOS latch (instead of a CMOS-based latch) to allow the structure function as a level converter during the regular operation. We multiplex the scan input and the TSV input. The select signal of this MUX (SCTRL) controls whether the TSV input or the scan input is applied to the input of the differential latch. Since this MUX needs to transfer the TSV voltage during TSV test, this MUX is designed using a transmission-gate-based MUX. The output of this MUX (IN_A) forms the one input of the differential latch. The second input of the differential latch (IN_B) is obtained by multiplexing the reference voltage and the inverse of IN_A. This MUX is designed as a tri-state-inverter-based MUX with V_{ref} as the supply voltage. During scan in or regular operation, the inverse of IN_A is connected to IN_B. During the TSV test, the signal TT is high, which ensures that IN_B is equal to V_{ref} . During scan-in or signal-recovery mode, TT is low, which ensures that the inverse of IN_A is applied to IN_B as shown in Table 1 and Fig. 18.

Table 1: Assignment of the control signals for the test circuit.

Control Signal Values in Different Mode					
Mode	SCTRL	SCLK	TT	EN	Comment
TSV Test	0	Scan CLK	1	0	EN=0 ensures logic inputs does not switch during TSV test or scan in to save power. After scan in EN is raised to high to apply the logic vectors
Scan IN/Out	1	Scan CLK	0	0	
Recovery	0	V_{DD}	0	1	EN is initially '0' to load the correct states in the NAND latch, then raised to '1' to ensure (a) NAND latch stores the correct states and (b) TSV or comparator outputs correctly applied to logic

Output MUX selection Logic						
INPUT			MUX Control Signals			Comments
EN	SCTRL	Q	S _{TSV} (select TSV directly)	S _{COMP} (select recovered signal)	S _Q (select scan output)	
0	X	X	0	0	0	
1	1	X	0	0	1	
1	0	0	0	1	0	
1	0	1	1	0	0	

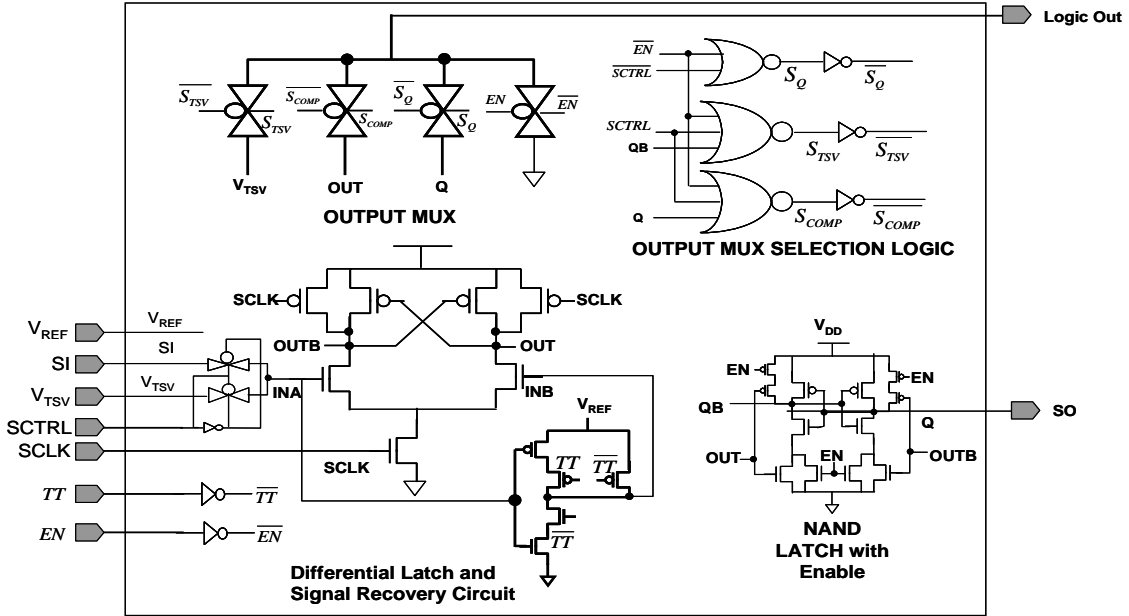


Fig. 18: Circuit schematic of the proposed test circuit.

While operating as a flip-flop during the TSV test and the scan-in/out mode, the enable signal (SCLK) of the differential latch is the scan-clock signal. However, for the signal recovery, SCLK is held high so that the circuit behaves as a level converter by multiplexing V_{DD} and the scan

clock. The selection can be achieved using TT and SCTRL. Note that the SCLK-generation circuit is a global one shared by all test circuits.

The output-selection logic is designed to ensure that the logic input is equal to: (a) the output of the scan flip-flop during pre-bond logic test, (b) the output of the comparator for TSVs requiring signal recovery in the operating mode, and (c) the output of the TSVs for good TSVs in the operating mode by multiplexing the three outputs. The control logic for the multiplexor is shown in Fig. 19.

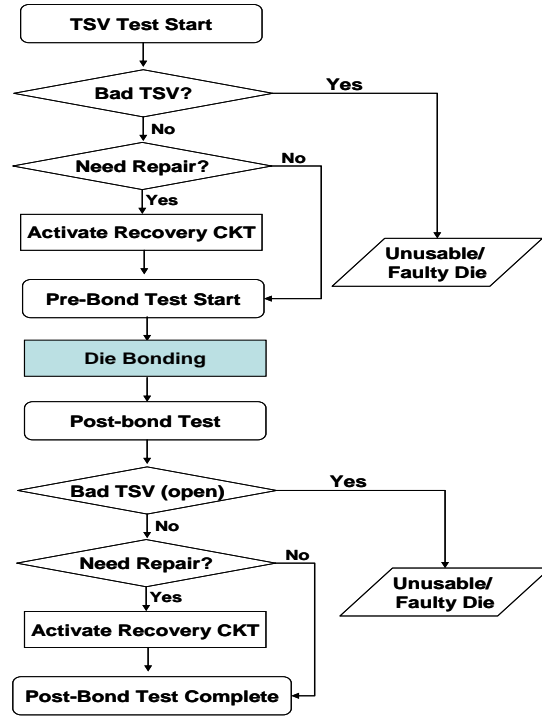


Fig. 19: The overall test philosophy.

Note that the select signal that differentiates between a scan mode and a regular mode is shared by all test circuits. However, during the operation, one needs to differentiate between direct TSV connections and comparator connection, which is a local signal. This selection is achieved by re-using the NAND-LATCH connected to the sense-amplifier-based flip-flop to store the local

requirements of the TSVs as shown in Fig. 19. To configure the NAND latches to the proper state, two options are explored. First, after pre-bond testing, the requirements for each TSV are stored in a ROM and loaded into the scan flip-flops before starting the operation. Second, the TSV test can be performed, while powering up the bonded 3D IC (as a built-in-self-test). The NAND latch is disabled in the recovery mode to ensure that the configuration information is not destroyed as shown in Fig. 18.

4.5 *Simulation Results*

This section presents simulation results to verify the functionality of the proposed circuit. Considering the presence of the differential pair, the circuit functionality is verified in commercially available IBM 90nm CMOS technology. This simulation allows verifying the circuit considering the built-in technology-characterized process variation model for devices.

4.5.1 Application to Pre-Bond TSV Tests and Signal Recovery

Verification of Functionality

Fig. 20 shows the waveform of the operation of the proposed circuit demonstrating (a, c) the V_{TSV} detection and (b, d) signal recovery. The proposed circuit can successfully detect the V_{TSV} value. In the recovery mode, the proposed circuit can successfully recover the low voltage swing.

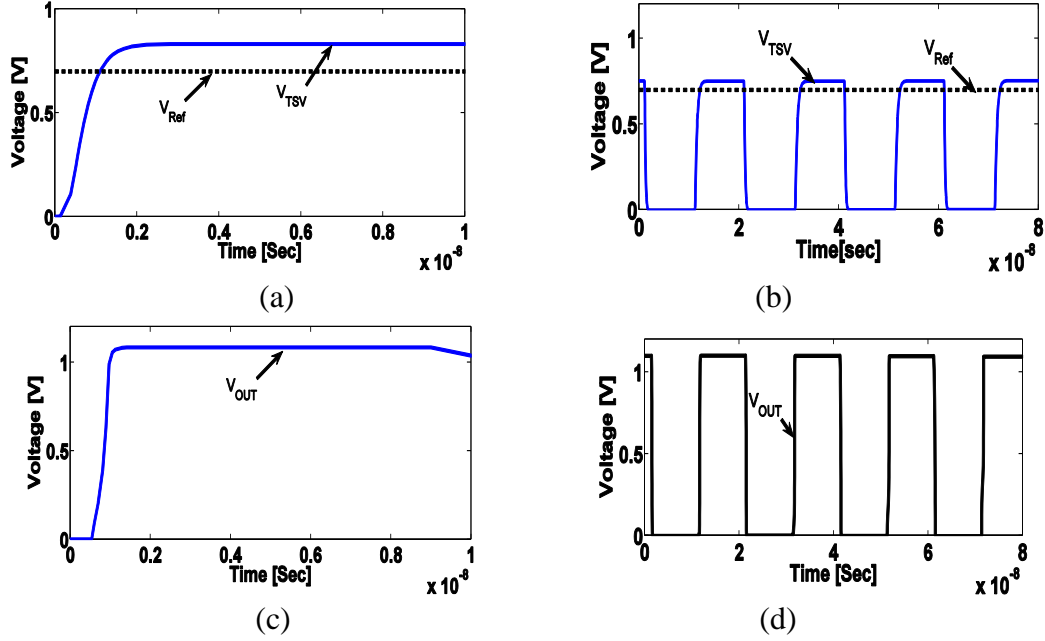


Fig. 20: Waveform of operation: (a) detection input, (b) recovery input, (c) detection output, and (d) recovery output.

We next consider statistical simulation of variation in the diameter of the short. A log-normal variation in the short diameter is considered, as shown in Fig. 21. The resistance corresponding to different short diameters is computed considering copper TSVs. As expected, the variation in the short diameter results in a variation in the resistance of TSV shorts. The resistance variation results in a variation in the voltage at the TTI-TSV connection (i.e., V_{TSV}). The proposed test structure successfully detects whether the short corresponds to a bad, repairable, or good TSV.

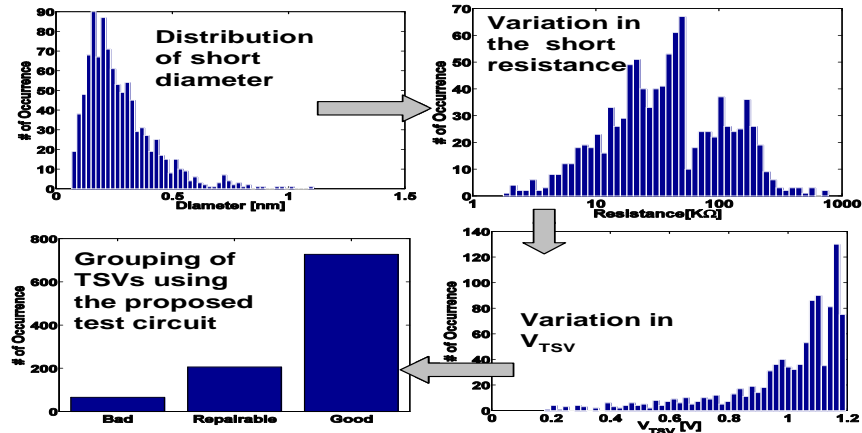


Fig. 21: Statistical simulation results showing the functionality during the pre-bond test.

Factors Affecting Detection Accuracy

Process Variations: A critical factor in the detection accuracy of the proposed structure is the effect of process variations. V_{TSV} for the same TSV short resistance varies because of the process variations in Fig. 22(a). The proposed design is optimized to reduce the offset by proper device sizing. The SPICE Monte-Carlo simulation is performed considering the internal variability model for the IBM 90nm technology [30]. The variability model simultaneously considers variations in all process parameters (such as L, W, and V_{th}). The process variation simulations illustrate the presence of two sources of detection error. Because of the variation in the strength of the PMOS transistor in the TTI, the generated V_{TSV} for different short resistances can also vary. The variation in V_{TSV} adds to the variation in the offset voltage of the differential latch because of device and output load mismatch. Considering these variations, the misdetection probability is computed. The random variations are considered in the short diameters for TSVs as shown in Fig. 22. For each TSV case, 1000 Monte-Carlo simulations are performed considering the random process variations, and whether the TSV is detected as the good, repairable, or bad ones is monitored. We compute the total probability of detecting the TSV as good, repairable, or bad and plot it against the short resistance. For an ideal case, the detection probability is a step function. However, because of finite offset, a misdetection probability exists, but it will be within an acceptable limit as shown in Fig. 22(b). Next, with the various standard deviation of the offset distribution, 1000 instances in a 3D die with ~1500 TSVs are considered. For each such instance, the TSV short diameter (and resistance) is randomly assigned, and the detection is performed. The total numbers of TSVs that are mis-detected as repairable or good ones are computed for each die instance as a percentage of the total number of TSVs. The average of this value over the 1000 die instances are plotted in Fig. 22(c). The average

percentage of misdetection error in detecting the bad TSV as a repairable TSV or vice-versa is very low, while marginally higher for detecting good ones as a repairable TSV or vice-versa.

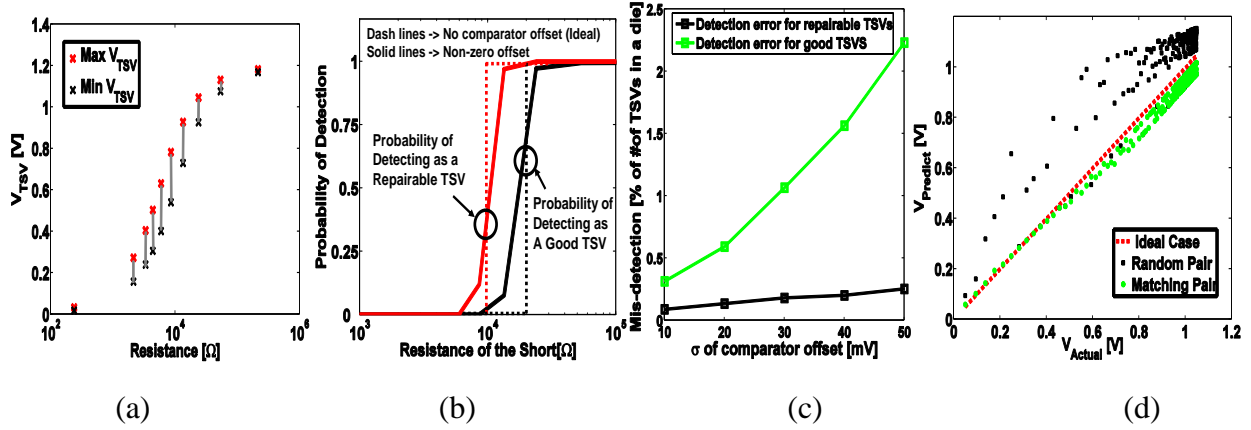


Fig. 22: Sources of detection inaccuracy obtained using Monte-Carlo simulations in 90nm CMOS: (a) V_{TSV} variation, (b) detection error, (c) effect of offset variation, and (d) driver-TTI mismatch.

Variation the Driver Strength: In a real 3D system, the drivers of each TSV are not identical. Hence, the extent of the signal degradation estimated with a fixed TTI (referred to as $V_{Predicted}$) may not exactly correlate with the actual signal degradation (referred to as V_{actual}). Fig. 22(d) shows a correlation with the actual signal degradation for different random driver sizes with a fixed TTI (referred to as a random pair). This simulation result is obtained through multiple Monte-Carlo simulations considering random variations in short resistance and process parameters. However, the size of the drivers of each TSV is known after the design and full-chip placement/routing of the 3D chip. We propose to use this information to match the size of the TTI for each TSV with the size of the actual driver of that TSV in the different die (or same die for output TSVs). As expected, such a matched pair significantly improves the correlation between the predicted and actual signal degradation. The marginal difference can still exist because of the random process variations as shown in Fig. 22(d). The above observation clearly

indicates that physical-design-aware synthesis of the TTI will help to improve the detection accuracy.

Signal Recovery and TSV-limited Functional Yield: This section addresses the effectiveness of the pre-bond TSV test and signal recovery on the yield enhancement as shown in Fig. 23. The offset distribution obtained from SPICE simulation is considered. Using the method mentioned in the previous section, 1000 random instances of a 3D die with 1500 TSVs and different short resistances are generated. The TSVs are grouped using the proposed circuit in bad/repairable/good groups. If required, the signal-recovery circuit is activated. A die is considered as a faulty one if any TSVs with signal swing $< 1V$ (with or without signal recovery) exist. As expected, with an increase in the variance in the short diameter, the number of good TSVs reduces while that of repairable or bad TSVs increases in Fig. 23(a). The use of signal recovery improves the TSV yield and allows a circuit to function with less control in the TSV process as shown in Fig. 23(b).

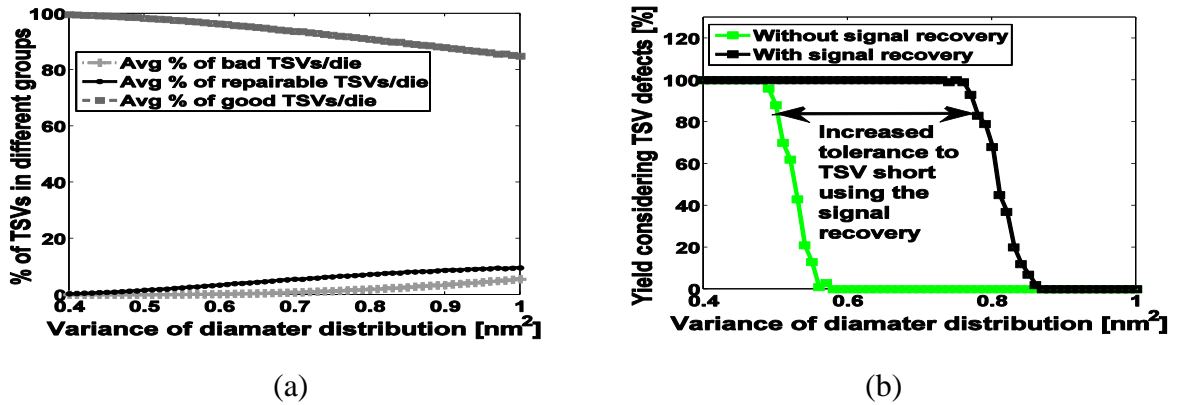


Fig. 23: (a) Detection of TSV groups and (b) effect of signal recovery on pre-bond yield.

4.5.2 Application to Post-Bond TSV Tests and Signal Recovery

This section studies the effectiveness of the proposed circuit in characterizing the variation in the post-bond resistance of the TSV to ensure good signal quality after bonding. Note that the post-

bond variation captures the combined effect of different sources such as weak open, misalignment, and variation in bonding resistance. After bonding, the test structure is activated to detect V_{TSV} . However, for the post-bond test, the signal driving inverter is in one tier, and V_{TSV} is sampled in the test circuit of the other tier. If a particular TSV under the test has appreciable R_{short} (i.e., a repairable TSV), the voltage level at the TSV output degrades. Note that signal-recovery circuit will be activated based on the outcome of the pre-bond test. The detection of post-bond resistance variation is particularly challenging for TSVs with no oxide short (i.e., “good” TSVs).

As in the pre-bond testing, the proposed circuit detects the signal quality even for good TSVs with no TSV short by activating NMOS of TTI as shown in Fig. 24(a). The statistical simulation is performed to verify the functionality of the post-bond test and recovery circuit as shown in Fig. 24(b). The log-normal distribution for post-bond TSV resistance is considered. As expected, variation in post-bond TSV resistance results in variation in V_{TSV} . Depending on the level of V_{TSV} , the proposed detection circuits successfully group TSVs as bad, repairable and good after bonding.

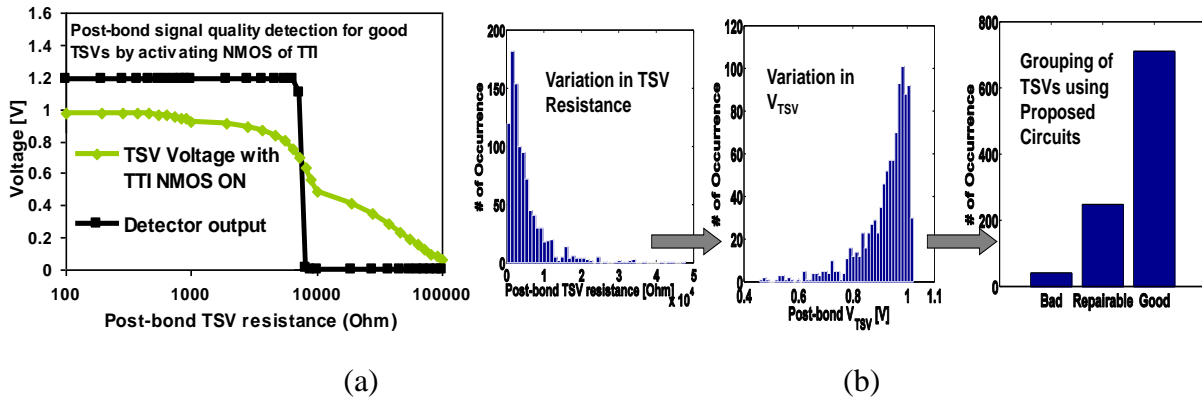


Fig. 24: Post-bond testing and recovery: (a) detection with the test circuit, (b) statistical simulation results showing the functionality at post-bond test.

Fig. 25 shows the effect of the variation in the post-bond TSV resistance on the signal swing and slew at the input of the receiving logic gate considering various oxide short resistances for good and repairable TSVs. The effectiveness of the signal recovery is analyzed to correct against post-bond resistance variation. As shown in Fig. 25(a) and 25(b), both the signal swing and the signal slew degrade significantly with a high variation in the post-bond TSV resistance. The effect is more pronounced with a less resistive short in the TSV. With the recovery, a small delay (compared to no recovery) through TSV for high post-bond resistance can be achieved. The delay improvement is primarily due to the improvement in signal slew (i.e., high signal slew increases the delay of the following logic gate). Note that the TSVs with very small variation in the resistance will bypass the signal-recovery circuit. For very small variation in the resistance, when the signal recovery circuit is bypassed, the slew at the input of the receiver degrades from ~14% of clock-high time to ~17% clock-high time. This degradation is primarily due to the resistance of the transmission gate at the output MUX.

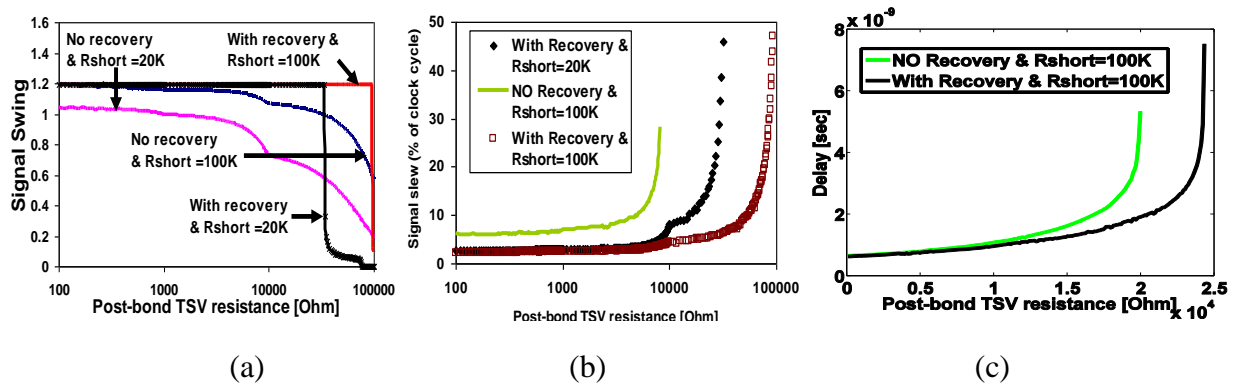


Fig. 25: Post-bond testing and recovery: (a) effect on signal swing, (b) signal slew, and (c) delay.

4.6 System-Level Full-Chip Analysis

This section analyzes the overhead of the proposed test structure considering the 3D full-chip physical design. For overhead analysis, the 45nm PDK technology model [31] and cell libraries

are considered. System-level full-chip analysis is performed by Chang Liu and Daehyun Kim in GTCAD Lab.

4.6.1 Target 3D Structure and Design Flow

FFT256_8 design [32] is used to demonstrate our experiment on a 3D circuit. The FFT256_8 is a design with 320K logic gates. The design is implemented in 45nm technology with 6 metal layers. The target 3D structure is shown in Fig. 26(a), in which the two dies are stacked in a face-to-back fashion with via-first TSVs. The TSV structure is shown in Fig. 26(a). The landing pad in M1 or M6 occupies 3 standard rows. A keep-out zone at the device layer, which occupies 4 standard rows, is considered in the simulation.

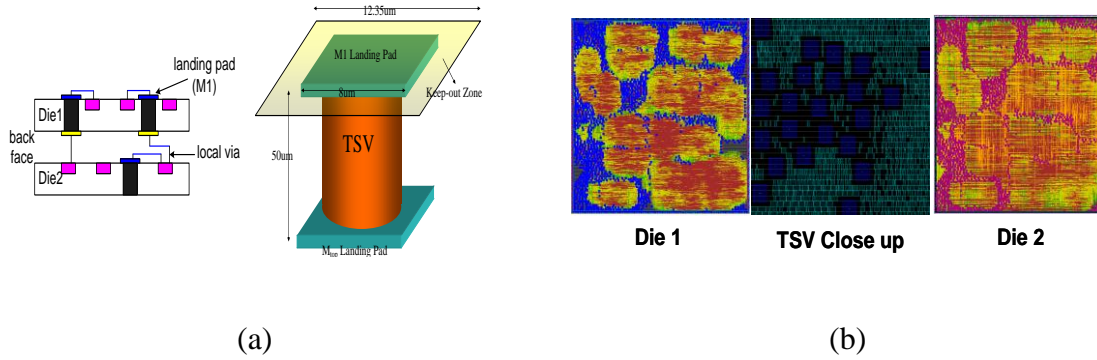


Fig. 26: The 3D system-level analysis: (a) the target 3D structure and TSV sizes and (b) full-chip layouts of the designed 3D stack for FFT256_8 circuit.

The 3D physical design flow [33, 34] is shown in Fig. 27. First, a min-cut practitioner is used to partition the top-level design into two dies. Each cut becomes a pin in each die, which corresponds to a TSV. Second, TSVs and standard cells sequentially are placed. The TSV pins are converted to TSV standard cells, which are defined in the physical library (.LEF). Third, the TSV cells and standard cells are placed together in the first die using the predefined pin locations as constraints. Finally, TSV standard cells are changed to back to TSV pins to do routing and optimization as the usual design flow. For the second die, the TSV landing pad locations from

the previous die are obtained. With these locations as constraints, the placement and routing are performed. The following steps are the same as in the first die.

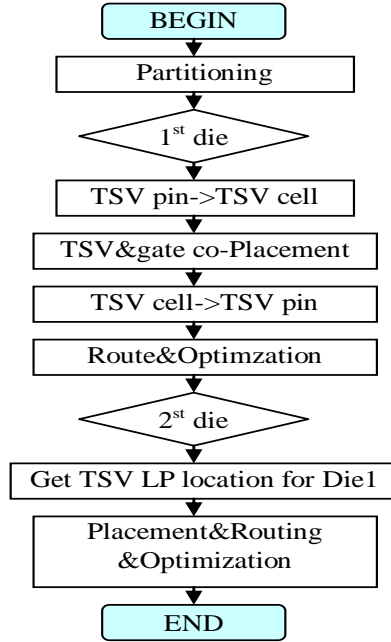


Fig. 27: The 3D design flow used in this work.

After all the designs are done, 3D timing analysis is performed using Primetime in combination of our own scripts. First, the top-level Verilog file containing these two dies needs to be generated. The top-level SPEF file for TSV parasitics is generated. After these files are ready, Primetime is used to read in the Verilog files and SPEF files in the incremental modes. The stitched SPEF file containing the RC information of both two dies and the TSVs is generated. Then, timing analysis can be performed on the stitched files. Fig. 26(b) shows the die shot of the two layers. Blue squares in Die 1 show the TSV M1 landing pads, and the pink squares in Die 2 show the TSV M6 landing pads. The total number of TSVs is 1444. The area per die is $1.08 \times 1.08 \text{mm}^2$.

4.6.2 Delay and Power Impact of TSV Test Circuits

This section considers the application of the proposed test circuit on the designed 3D system of the FFT256_8 circuit as shown in Fig. 26(b). The proposed test circuit is re-designed in 45nm PDK technology, and the functionality is verified. As the 90nm CMOS counterpart, the 45nm design also performs correct detection and recovery. The area, delay, and power dissipation of the proposed test circuit are estimated in 45nm node for inclusion into our 3D design flow. The physical area of the proposed design in 45nm technology is $\sim 21\mu\text{m}^2$. The additional cell is added to each TSV in our 3D design flow to estimate the area overhead. For the area overhead estimation, different partition options are considered, which results in different number of TSVs for the entire 3D chip. The area overhead of the proposed design was observed to be less than 4% of the total die area when TSV area is $\sim 20\%$ of the die area as shown in Fig. 28.

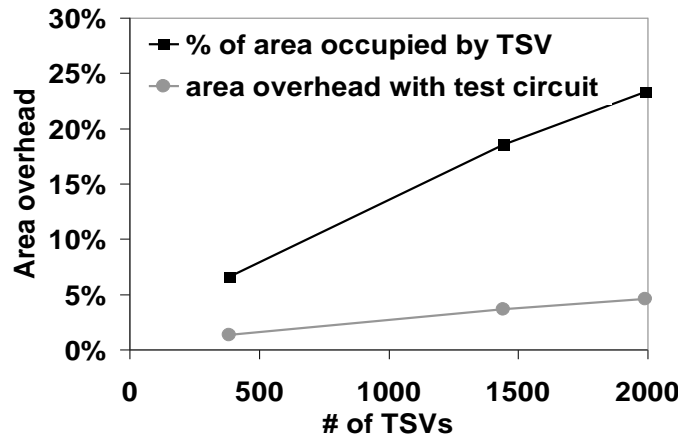


Fig. 28: The area overhead of the proposed structure from system-level analysis.

CHAPTER V

DESIGN METHOD TO CHARACTERIZE AND COMPENSATE FOR TEMPERATURE VARIATION IN MANY-CORE SYSTEMS

As the silicon device scales down, the dynamic power in successive generations reduces. However, the leakage power is expected to increase continuously. Because of an increased leakage power and device scaling, the power density of a chip increases significantly, which in turn increases a peak temperature. A high temperature degrades the reliability and the performance of systems. Beyond a threshold temperature, the systems enter thermal runaway condition, which creates permanent damage in the systems. The increasing variations in the manufacturing process will add to the increase in the leakage power leading to wide chip-to-chip leakage variations. The increasing chip leakage coupled with the increase in the number of cores in many-core systems can lead to unsustainable increase in a chip power and a die temperature imposing stringent challenges in the test and normal operation of many-core systems.

The proposed research explores a design methodology to handle temperature variation both during the test and normal operation in digital systems. To verify the proposed approach, many-core systems are used as an example of digital systems. Many-core systems have been attracted for decades because of the advantages of multi-functional complex operations and high performance. Although the proposed design methods primarily focus on many-core systems, the proposed research can be applied to other digital systems. As an example of test applications, burn-in test is considered.

5.1 *Burn-in Test*

Burn-in test is used to screen infant defects at the early stage [35]. Fig. 29(a) shows the number of

defects over the life time of a chip called a “Bathtub” curve [35, 36]. This curve indicates that the number of defects detected in the chip is relatively large during the “infancy” of the chip. Burn-in test is mainly designed to detect these defects by accelerating the aging of devices. Burn-in test places devices under test (DUT) at the 1.3~1.4X elevated temperature and voltage to accelerate the aging of devices. For example, Fig. 29(b) shows the effect of temperature on burn-in time based on the acceleration models from [36, 37]. The acceleration causes a left-shift in the bathtub curve indicating that the defects are caused to occur early so that the defects can be detected before shipping the chip to the customer. Various types of burn-in tests such as DC, dynamic, monitored, and test-in burn-in (TIBI) are performed depending on the application and required test coverage. In this work, the dynamic burn-in test is considered. In the dynamic burn-in test, along with the raised supply and temperature, input vectors are also applied to the inputs, but the outputs are not monitored. The purpose of applying test vectors to the inputs is to toggle the internal nodes of the chip. Since the cost of the burn-in test is a significant fraction of the overall budget, the test cost and yield loss during burn-in test need to decrease [38]. However, the test time, test cost, and yield loss reduction need to be achieved without sacrificing the test coverage. Significant efforts have been directed to reduce the burn-in time with reasonable fault coverage for single or few core processors [39-41]. With the advent of the many-core chips, understanding and addressing the burn-in challenges for many-core chips become an important problem.

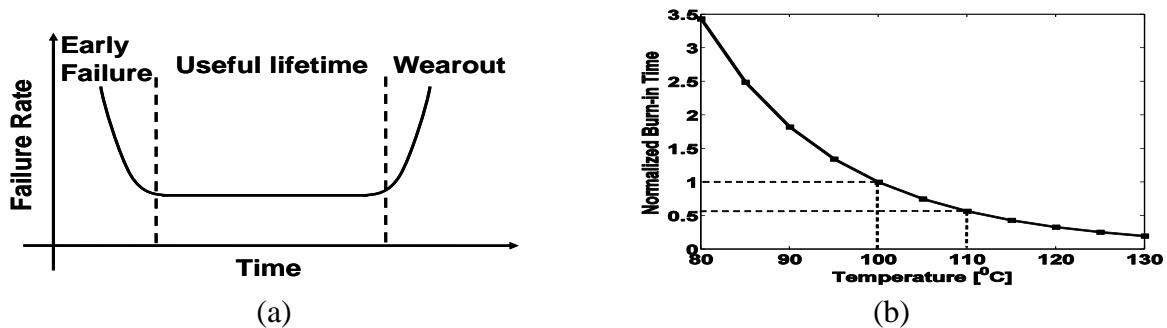


Fig. 29. Burn-in basics: (a) bathtub curve and (b) effect of temperature on acceleration.

5.2 *Challenges of Burn-in Test for Many-Core System*

Leakage Variation and Leakage-Temperature Interaction

Both during wafer-level and package-level burn-in test, the ambient temperatures are maintained by the burn-in chamber. The silicon temperature (T_{si}) depends on self-heating effect in the die (i.e., the power dissipation in the chip and the thermal resistance between silicon and ambient). For the same thermal resistance, an increase in the leakage power and total power can result in different silicon temperature. The die-to-die variation in the silicon temperature reduces the test quality [42, 43]. Moreover, as the leakage power increases exponentially with temperature, the interaction builds a positive feedback system, which may cause a thermal runaway without a careful control. Because of the leakage-temperature interaction, in a many-core chip, the total power becomes a super-linear function of the number of cores being simultaneously stressed. Hence, the controllability of the silicon temperature imposes a stringent requirement on the number of cores that can be stressed simultaneously.

Power Delivery during Burn-in Test

Because of the 1.3X higher stress voltage (i.e., high dynamic and leakage power) and high temperature, each core consumes significantly high power (i.e., draws high peak current) in burn-in test than the nominal condition. Because of high power, the power network may not be able to deliver full power to all cores during burn-in test as the IR drop and Ldi/dt noise in the delivery network is high. Even if the on-chip power regulator can address the IR drop challenge, the high Ldi/dt becomes an unsustainable problem (i.e., a high V_{max} can cause permanent damage to the device, while low V_{min} and high supply settling time degrade the quality of test).

Spatiotemporal Non-uniformity in the Thermal Field

During the burn-in test, all the cores are kept at a constant temperature to obtain the proper defect coverage. The stringent requirement on the number of simultaneously stressed core (NSSC) forces an undesirable non-uniformity in the thermal field. The temperature of the stressed cores becomes much higher than the unstressed ones, which create highly non-uniform on-chip spatial thermal field. The nature of this thermal field varies over time. This spatiotemporal non-uniformity reduces the confidence in the exact silicon temperature, which in turn degrades the test quality.

Chip-to-Chip Variation in Test Quality and Burn-in Time

Since all cores cannot be simultaneously stressed, the total burn-in time for the chip becomes high. If the temperature rises at a fast rate, the number of simultaneously stressed cores needs to be reduced. Further, because of the chip-to-chip leakage variations [44, 45], the total chip power for a fixed NSSC also varies from chip to chip. To ensure that thermal runaway does not occur even for the low-Vt dies (i.e., high leakage power), the number of simultaneously stressed cores needs to be very low. However, the chip-to-chip variation will result in a lower temperature than the required silicon temperature for high-Vt dies (i.e., low leakage power, $T_{si} < T_{stress}$) reducing the quality of test. Further, the burn-in time will be unnecessarily long for the high-Vt dies. On the other hand, if the number of cores is decided based on the high-Vt corners, the thermal runaway can occur for the low-Vt corners. In summary, the average burn-in time can be very high to ensure minimal yield loss because of thermal runaway.

5.3 Modeling and Simulation Framework

To demonstrate the effectiveness of the proposed approach, a tile-type many-core architecture, in which a set of tiles (cores) is arranged in two dimensions, is considered [47, 48]. Each core

consists of a logic core (simple cores with 10 million gates operating at 3GHz) and a local cache.

The cores are connected by a mesh network. Fig. 30 shows the overall modeling framework.

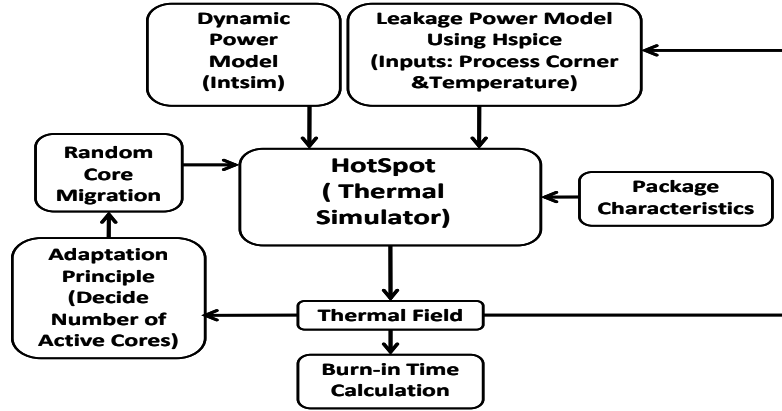


Fig. 30: Modeling framework.

5.3.1 Power Models for Cores

Leakage Power Models: A critical path is considered length of 10 (i.e., 10-stage 2-input NAND gate with a fan-out of 4). The sizes of gates in this chain are determined to meet the target frequency of 3 GHz at nominal supply voltage. As mentioned earlier, the sleep transistor is sized to have 5% delay penalty for the 2-input NAND gate-based critical path. The leakage power for a 2-input NAND gate with designed size was calculated considering the average of all possible input vectors. The leakage is computed using circuit simulations with the predictive technology models. The Monte-Carlo simulation considering Gaussian threshold voltage distributions is performed to model the effect of process variation on the leakage power. A set of leakage-temperature interaction models for the 10 million 2-input NAND gate circuit is generated considering “on” and “off” sleep transistor and different process corners.

Dynamic Power Models: The dynamic power of the core consists of the switching power of the internal nodes, interconnect power, and clock power. The technology-driven interconnect simulation tool (IntSim) proposed by Sekar et. al [49] is used to estimate the interconnect and

clock power. The each logic core is assumed to be 0.75mm x 1.5mm to compute the average interconnect power and clock power using the IntSim. The node-switching power is estimated from circuit simulations.

Transition Power Models: During a power-migration event, the supply rails of the stressed and unstressed cores need to be discharged and charged respectively, which results in additional transition power. The capacitance of the supply rail of this predictive core is computed to estimate this capacitance for both transition power and supply noise computation. The junction capacitance of the 20 million gates (for 10 million 2-input NAND gates) is first estimated. Next, a voltage grid is assumed for the each core designed with at the global metal layer. The resistance, substrate capacitance, and inductance of this supply line are estimated from the predictive interconnect models [31]. The total capacitance of the virtual supply node is computed by adding this metal capacitance, the PMOS junction capacitance, and 10% additional core-level decoupling capacitance.

5.3.2 Thermal Modeling Framework

The dynamic variations in the locations of the stressed cores create the spatiotemporal variations in the full-chip power profile, which is connected to the distributed RC-based thermal simulator (Hotspot) [51]. The transient thermal simulations are performed considering the ambient temperature of burn-in condition. The temperature-leakage interaction curves for different process corners and the operation of a sleep transistor are considered in the simulation. The coupled simulation helps more accurate characterization of the impact of process corners and thermal resistance to ambient on the on-chip thermal profile. As the power migrations are performed after each time-slice interval, the power profile of the chip is internally modified to account for the additional transition energy during migration. To change the NSSC, the

temperature values obtained from the sensors are checked. The sensor locations on the chip are predefined, and these sensors have a finite sampling interval.

5.3.3 Modeling of Power Supply Noise

To model the effect of NSSC on the power supply noise, a distributed RLC-based model of the supply network is created. The RLC values of the core-to-core supply grid are estimated. The current variations in each clock cycle are assumed to be triangular. To characterize the supply noise, we compute two critical parameters: (a) the minimum and maximum supply noise and (b) the time required for supply noise to stabilize.

5.4 *Adaptive Spatiotemporal Power Migration (ASTPM) Architecture*

The proposed burn-in test architecture with adaptive spatiotemporal power migration (ASTPM) is based on the two fundamental principles: (a) spatiotemporal power migration and (b) adaptation of the NSSC depending on the thermal field. This chapter first explains the above principles and next presents the overall architecture of ASTPM.

5.4.1 Spatiotemporal Power Migration (STPM)

The thermal field of a many-core chip depends on the three parameters: namely, heat generation in the chip, heat outflow from the chip, and heat redistribution within the chip. During burn-in test, the total power dissipation in the chip is related to the supply voltage and required temperature. Hence, reducing the power dissipation requires either the reduction of supply voltage or reduction in the NSSC. Clearly, the first one is not a feasible option as it will exponentially decrease the acceleration factor. The second one is a more feasible option, which will be used for adaptation. For the time being, consider a fixed NSSC. For fast burn-in test, one would like to maintain a high NSSC. However, a high NSSC will imply a fast rise in temperature

and the possibility of thermal runaway. Therefore, our aim is to reduce the maximum temperature and the rate of increase in temperature for a fixed NSSC (i.e., constant heat generation) and given cooling solution, which will automatically translate to a high number of NSSC for a given target burn-in temperature.

The proposed research achieves this goal by continuously redistributing the generated heat in space and time. In a many-core chip, heat redistribution can be achieved by spatiotemporally varying the location of the heat generation points (i.e., stressed cores) at a time interval (i.e., migration interval) smaller than the time constant of the temperature rise as shown in Fig. 31. This approach is referred to as spatiotemporal power migration (STPM). STPM is performed continuously for all of the stressed cores and for any NSSC instead of migrating cores individually only when their local peak temperature constraint is violated. Therefore, it is a coordinated and proactive approach. The unstressed cores or “off” cores are clock-gated and supply-gated to dissipate minimal power, while an “on” or stressed core receives the full stress voltage and dynamic activity.

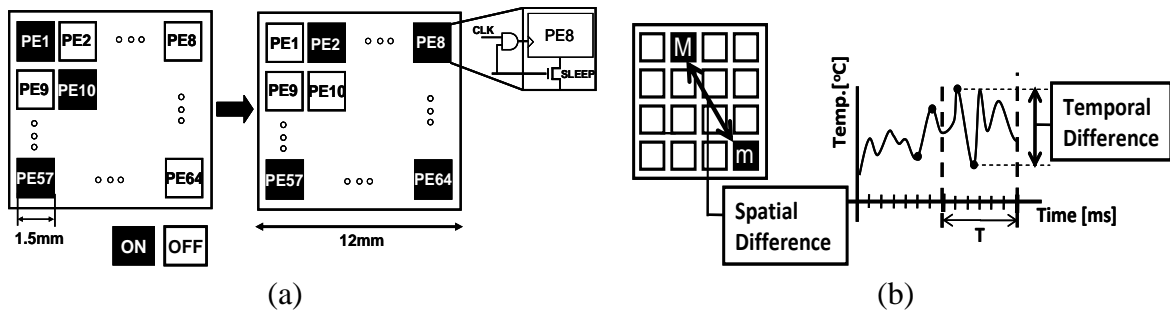


Fig. 31: (a) Concepts of spatiotemporal power migration and (b) spatial and temporal difference.

The effect of power migration can be explained through the thermal RC behavior. The temperature at a location is analogous to voltage across the thermal capacitance at that location. The heat generation is analogous to a current source with the magnitude of the power dissipation. The longer the current source at a location is active (i.e., core is stressed), the higher the voltage

(i.e., temperature) at that location. If the power migration occurs more often, the current sources at specific locations are “on” for a smaller time length, and the capacitor charges to a smaller voltage level. A slow rise of temperature at a location implies a low maximum temperature. Hence, STPM will result in a low maximum chip temperature for given number of the NSSC. For the target burn-in temperature, the available thermal slack can be used to stress more cores at the same time. STPM also ensures low temperature variation at a location over time and the reduced temperature difference between a “on” and “off” location. Both of the above factors imply that all the cores experience a similar accelerating factor improving defect coverage and the quality of test.

5.4.2 Adaptive Spatiotemporal Power Migration

As explained earlier, for a given temperature target and time-slice interval, STPM allows increasing the NSSC. However, the maximum value of NSSC that can be used without violating thermal runaway condition depends on the process corner and thermal resistance to ambient. For example, for the low-V_t dies, because of the high leakage power, the total power of a stressed core is high. On the other hand, low leakage in the high-V_t corners provides the opportunity for increasing NSSC. Similarly, a low thermal resistance to ambient (i.e., improved heat outflow) will allow high power to be dissipated (i.e., higher NSSC). Therefore, using a predefined NSSC for every die and burn-in condition is not efficient as it will result in die-to-die variation in the silicon temperature degrading the quality of test. Using a NSSC suitable for nominal process corners, while testing a low-V_t die, may cause permanent damage and thermal runaway. On the other hand, using the same NSSC will reduce the silicon temperature for high-V_t dies. Use of less than maximum allowable NSSC will waste burn-in chamber time.

The proposed approach adapts the NSSC during burn-in test depending on the maximum chip temperature to address the above mentioned challenge. This maximum temperature in the chip is sampled by on-chip sensors in the cores. If the maximum temperature in the chip crosses a specific threshold, the controller starts to reduce the number of active cores. Similarly, if the maximum temperature in the chip drops below a certain threshold, the controller increases the number of active cores. Consequently, the NSSC stabilizes to the optimal value, which will ensure the constant silicon temperature for all dies and a protection against thermal runaway.

If the on-chip process sensors are available, the adaptive method can also be used to reduce burn-in chamber time. The maximum NSSC for a process corner is strongly correlated to the chip leakage. The proposed method senses the leakage of the chip and divides the wafers (or dies for package-level burn-in) into different process bins (e.g., 3 bins – high-Vt, nominal-Vt, and low-Vt). The maximum leakage for a particular bin is used to estimate the NSSC and the burn-in time for all the wafers (or dies) in that bin. Since the maximum leakage in the high-Vt bin is lower than that of the low-Vt bin, the dies in the high-Vt bin will be tested with higher NSSC. Hence, a low burn-in chamber time will be sufficient compared to those in the low-Vt bins.

5.4.3 Overall Test Methodology

Fig. 32 explains the overall methodology of the proposed ASTPM-based burn-in test. If process sensors are available, the proposed method first detects the process corner to select the expected number of NSSC and burn-in time. Otherwise, a starting point of a NSSC is selected (~50% of the total cores), and the maximum burn-in time is considered. After the warm-up period, as the silicon temperature reaches close to the target value, the stress is applied. During the test, the location of the stressed cores is varied randomly after each time-slice interval using STPM principles. Note that random migration (i.e., a new set of stressed cores is selected randomly at

the end of each time-slice interval) policy is considered. However, more sophisticated migration policies can also be considered. The sensors sample the temperature value at finite sampling intervals. The sensor-sampling interval can be higher than or equal to the time-slice interval. If the maximum temperature of the chip is higher (or lower) than the required value, the NSSC is increased (or decreased) and the STPM continues. A counter tracks the “on” or “off” states of every core. After the stress time for all cores reaches the burn-in time, the test is completed.

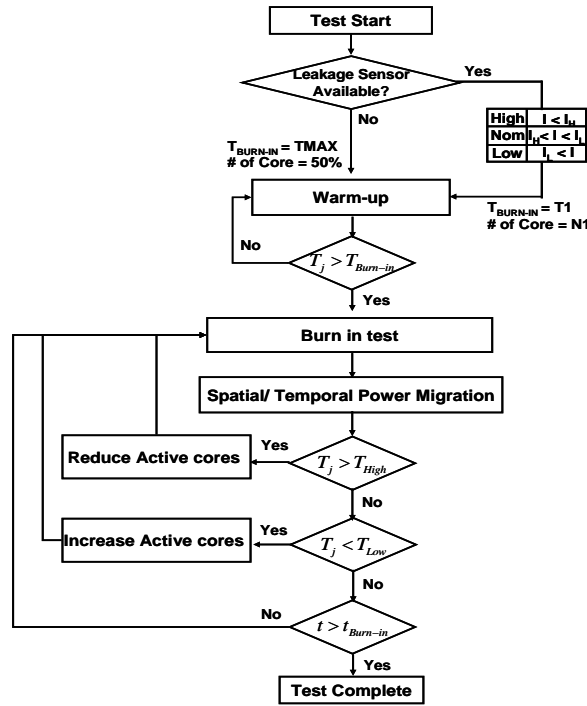


Fig. 32: Adaptive spatiotemporal power migration.

5.5 Simulation Results

This section presents the simulation results for many-core burn-in test. First, the results illustrating the challenges in the many-core burn-in test are presented. Next, the effect of spatiotemporal power migration is presented, and finally the effect of adaptation of number of cores is shown.

5.5.1 Challenges in Many-Core Burn-in Test

Leakage-Temperature Interaction and Controllability of Junction Temperature across Process Corners

The leakage and temperature create a positive feedback loop, which can lead to significant on-chip temperature rise and thermal runaway. The strength of this interaction depends on the process corners. At low-Vt corners, leakage increases at a fast rate with temperature as shown in Fig 33. We consider burn-in test for different process corners without a leakage control for the 64-core chip with all cores being simultaneously tested. Fig. 34 shows low-Vt corner reaches high temperature (145°C) compared to other process corners. With a fixed number of NSSC, the silicon temperature will significantly vary from one chip to another chip. Additionally, significant temperature rise even in nominal and high-Vt corners is observed. This result suggests that the leakage-temperature interaction will limit the NSSC.

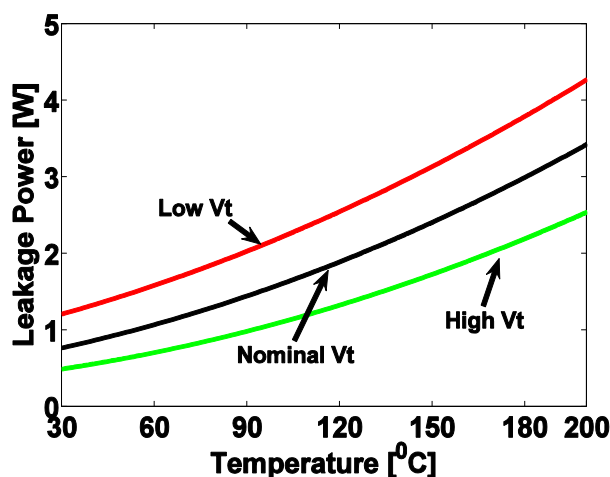


Fig. 33: Effect of temperature on leakage with process variations.

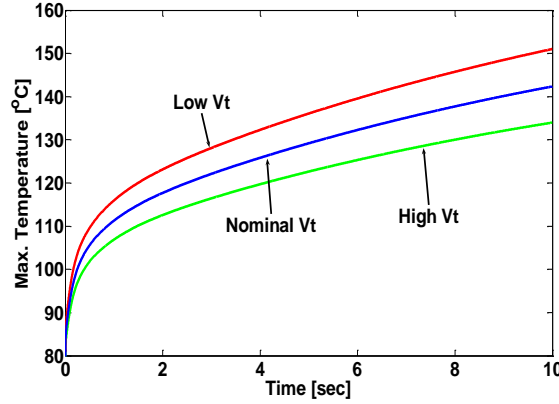


Fig. 34: Self-consistent leakage-temperature simulations.

Power Delivery Limits

In principle, for thermal control, one can use throttling (i.e., keep NSSC equal to 64). When temperature crosses the target value, all cores are simultaneously deactivated. However, such an approach imposes significant challenge on the power delivery. First, the IR drop on the delivery network will be high. Second, the Ldi/dt noise occurs during the activation and deactivation of cores. To understand the effect, IR drop and Ldi/dt simulations with different number of NSSC are performed as shown in Fig. 35. The resistance of the power-delivery network is designed to keep IR drop under 50mV when all cores are active and dissipate the nominal power. The resulting width of the metal lines is used to compute the metal inductance and capacitance. As the number of cores being simultaneously activated or deactivated increases, the increase in the total current results in (a) high 1st voltage droop (V_{min}), (b) high 1st voltage overshoot (V_{max}), and (c) long supply settling time (voltage variation is within ± 50 mV of stable supply). The low V_{min} degrades the quality of test, and the high voltage overshoot can permanently damage the devices. The long settling time increases the test time. Based on the experiment, all cores cannot be activated or deactivated simultaneously. If the NSSC is limited to 50% of total cores, the Ldi/dt noise significantly reduces.

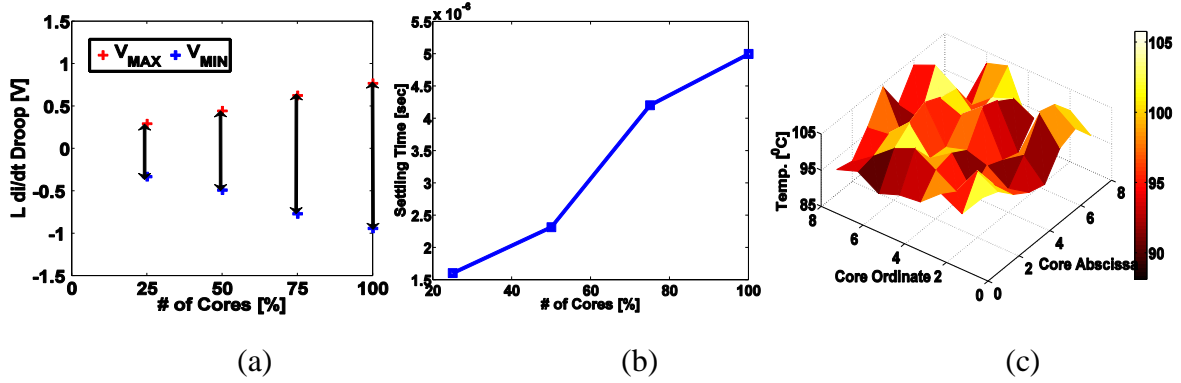


Fig. 35: Burn-in challenges for many-core: (a) Ldi/dt droop, (b) supply settling time, and (c) non-uniformity in the thermal profile.

Spatiotemporal Non-uniformity in the Thermal Field

With the NSSC limited to 50%, thermal simulation is performed. A set of 32 cores is kept active until the peak temperature violates the target constrain. Fig. 35(c) shows the spatial thermal field can have significant non-uniformity, which suggests that large variation in the stress applied to the devices across the chip degrades the quality of test. Further, the thermal field varies significantly over time depending on the location of stressed cores.

5.5.2 Spatiotemporal Power Migration

Even with a subset of active cores on, the temperature during burn-in test is difficult to control. Fig. 36 shows a fast rise in temperature for both a random assignment of active cores and “the best possible” assignment of on-cores. The best possible assignment corresponds to a checkerboard-type assignment of the on-cores. A better spatial assignment can more effectively distribute the generated heat in space and exploit the lateral heat flow in silicon. However, it cannot exploit the finite thermal capacity of silicon. The proposed spatiotemporal power migration method can simultaneously redistribute the generated heat in space and time. For the STPM, the locations of the stressed cores are changed continuously in every time-slice interval. Fig. 36 shows the effect of STPM on a fixed NSSC. The temperature rise is significantly slow, which leads to a low maximum temperature.

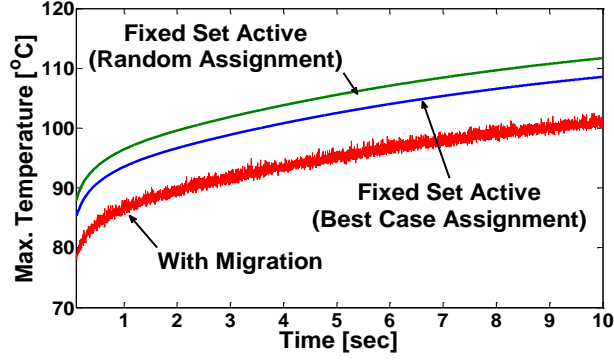


Fig. 36: Temperature rise with and without STPM.

Effect of Migration Interval

The migration interval (i.e., time-slice interval) has a strong impact on the thermal behavior of the system. Large time interval has a high maximum temperature as shown in Fig. 37(a) and 37(b). Similarly, if the migration interval is small, the fluctuation of the maximum temperature is also small (i.e., a tight control of temperature). Fast migration also significantly reduces the spatiotemporal non-uniformity as shown in Fig. 37(c) and 37(d). The on-chip spatial temperature variation (i.e., spatial difference) is low with fast migration. Similarly, the both the average value and spread of the core-to-core variations in the temporal difference are also significantly reduced with fast migration. This narrow distribution of temperature across cores over time and space implies more uniform temperature stress for all cores (i.e., better test quality). However, reducing the migration interval beyond a certain limit will incur transition time and energy penalties. Additionally, whenever a set of cores is activated and deactivated, a finite time interval is required for the supply to settle down.

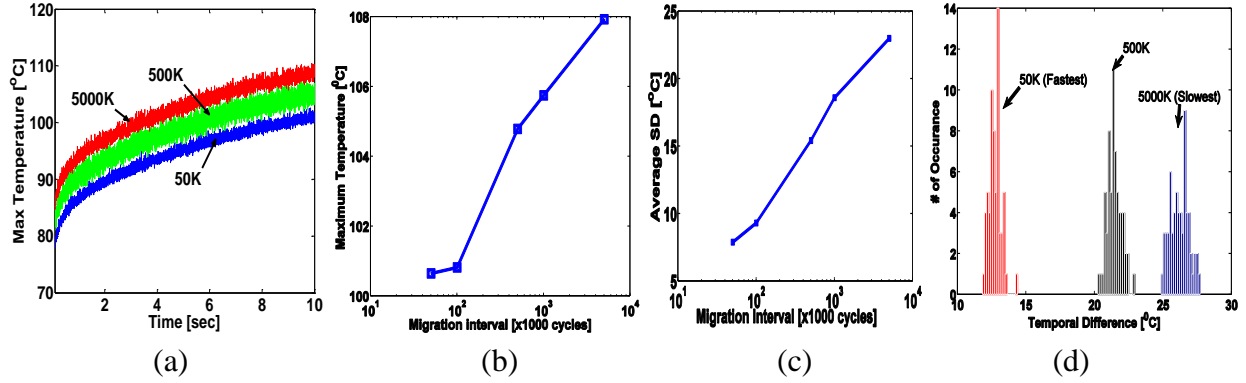


Fig. 37: The effect of migration interval on (a) max temp. vs. time, (b) max temp. vs. number of active cores at 6 sec, (c) spatial difference vs. time interval, and (d) temporal difference vs. time interval.

Effect of Number of Simultaneously Stressed Cores

Fig. 38 shows the variation of the maximum temperature over time for different number of simultaneously stressed cores (NSSC) for a constant migration interval. Fig. 38(a) shows that the maximum temperature increases with a high NSSC although the rate of increase is small with STPM. If the NSSC is high, the temperature fluctuation is reduced as shown in Fig. 38. Hence, if thermal limits allow, a high NSSC is highly beneficial as it: (a) leads to low burn-in time and (b) improves the quality of test. The above observation suggests that a high number of cores can be simultaneously stressed. Therefore, if applied, STPM with fast migration can improve test time and the quality of test.

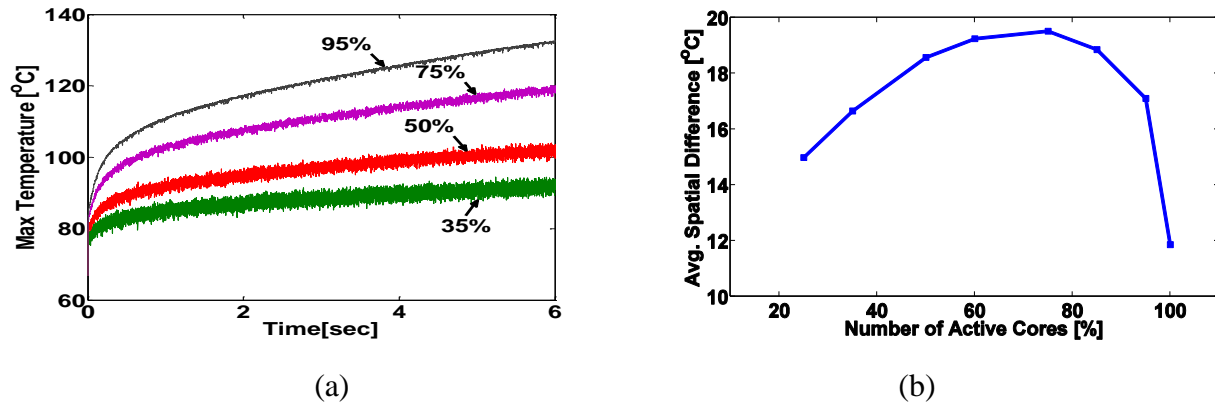


Fig. 38: The effect of number of active cores on (a) max temperature and (b) spatial difference.

5.5.3 Need for Adaptive Spatiotemporal Power Migration

The preceding discussion shows that with different migration interval, the allowed NSSC is different. Therefore, *it is imperative that NSSC needs to be adapted based on the migration interval*. However, for a fixed NSSC and migration interval, the process corner has a strong impact on the thermal profile. The Normal distribution of die-to-die threshold voltage is considered, which results in significant variation in the leakage current as shown in Fig. 39(a). Therefore, with the same NSSC (50% of total cores) and migration interval (100,000 clock cycle), the temperature of low-Vt corner increases fast and reaches to high value as shown in Fig. 39(b). In the high-Vt corner, temperature increases slowly and reaches a low max temperature. The on-chip spatial difference in temperature is also high for low-Vt corners as shown in Fig. 39(c). The difference is more pronounced for large migration interval and high NSSC. Hence, to keep the maximum temperature at a target level, a high NSSC is required for high-Vt corners while a low NSSC will require for low-Vt corners. These results show the need for *ASTPM* to control the die-to-die variation in silicon temperature and thermal non-uniformity.

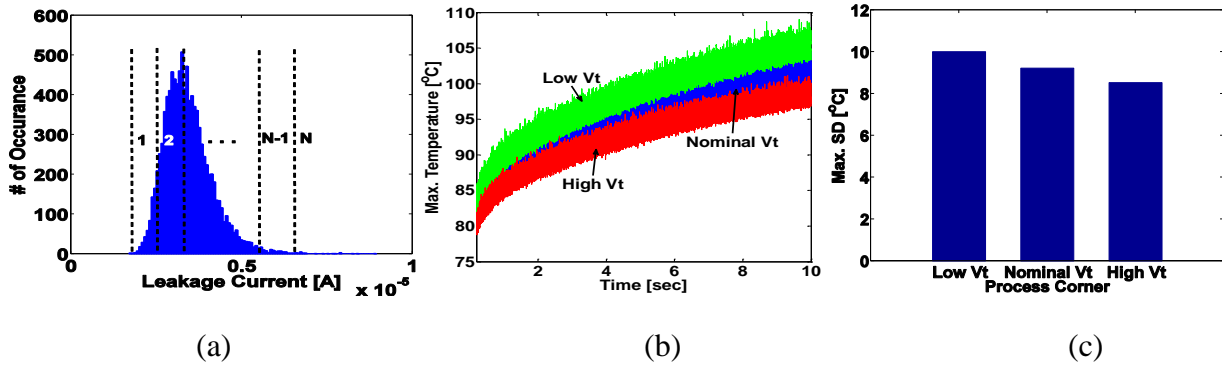


Fig. 39: The impact of process variation on (a) leakage power distribution, (b) max temperature, and (c) spatial difference.

5.5.4 Adaptive Spatiotemporal Power Migration

The proposed ASTPM monitors the maximum temperature and adaptively changes the NSSC based on that temperature. The effect of ASTPM on the maximum temperature for low-Vt, nominal-Vt, and high-Vt corners is shown in Fig. 39. The burn-in temperature needs to be in the range of 100-110°C. The observations on ASTPM are summarized below:

- Temperature control over time: For the nominal-Vt corner, the maximum temperature always remains in the target range.
- Reduced die-to-die variation in the silicon temperature: Unlike Fig. 39 (the fixed NSSC), the maximum temperature remains in the desired range for all the corners with the ASTPM. The NSSC initially changes with time and finally reaches a steady state.
- Effect on the NSSC at different corners: Fig. 40(a) shows the steady-state NSSC values for three corners. As expected, high-Vt corner allows higher NSSC than low-Vt corners.
- Effect on die-to-die variation on non-uniformity: The on-chip core-to-core temperature variation (i.e., spatial difference) is measured. The core-to-core temporal difference after the steady state is reached for every corner as shown in Fig. 40(b) and 40(c). The core-to-core temperature variations are very similar for all three corners with adaptation as shown in Fig. 40(b). The temporal difference is also similar as shown in Fig. 40(c).
- Effect of migration interval: ASTPM adapts the NSSC to its optimal value for different migration interval as well. As expected, fast migration allows high NSSC. ASTPM adapts the NSSC across the process corners for all migration intervals as shown in Fig. 40(a).

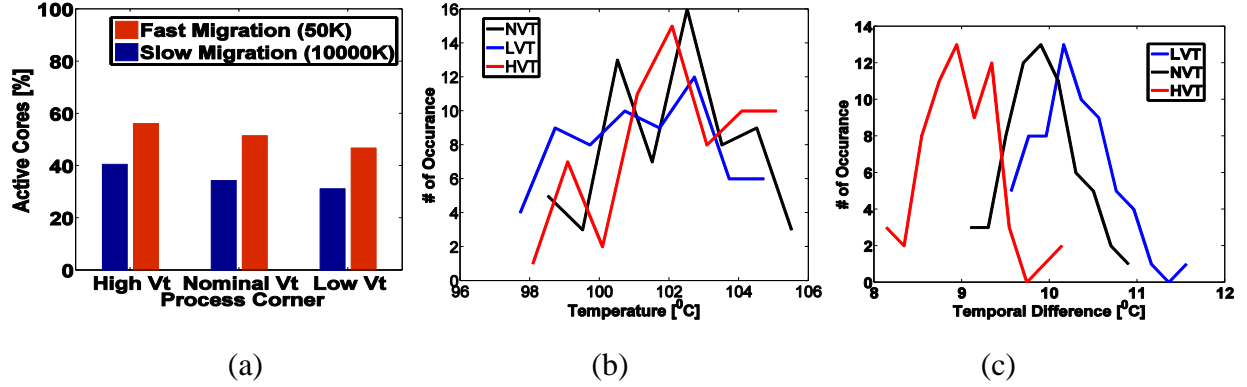


Fig. 40: Effect of ASTPM on (a) number of active cores, (b) core-to-core temperature variations, and (c) core-to-core temporal difference variations.

5.6 Burn-in Time Estimation

The acceleration factor during the test determines the burn-in time and the reliability coverage. For simplicity, the acceleration factor for the gate-oxide reliability is considered based on [37, 39]. The overall acceleration factor is the product of voltage (A_V) and temperature (A_T) acceleration. The well-known exponential dependence of the acceleration factor on voltage and temperature is used in the computation. A high voltage and temperature increase the acceleration factor and reduce the time to complete the burn-in test.

In single-core systems, in which all devices are stressed simultaneously at constant temperature and voltage, acceleration factors are same. However, in many-core systems, to avoid thermal runaway and maintain supply noise reliability, all the cores cannot be stressed simultaneously for the extended period of time. In the proposed ASTPM scheme, the locations of the stressed cores are varied randomly over time. Hence, for a given time interval, a core is stressed if and only if a core receives the voltage stress and clock signals (i.e., only when core is on). The proposed research assumes that the acceleration factor is negligible when voltage stress is very low (i.e., when sleep transistors are off). Therefore, the average acceleration factor for a core over a time period (T) will depend on how long the core was on. To estimate this factor for

a core A_j , the time period (T) is divided into several finite time intervals (t_i). For each interval, the acceleration factor for the core is computed based on the average temperature in that interval and the stress voltage. If the core is off, the acceleration factor is assumed to be “0”. The average acceleration factor for the core $[A_j(T)]$ is computed as

$$A_j(T) = \frac{1}{T} \times \sum_{t_i=0}^T A_T(t_i) \times A_V(t_i). \quad (14)$$

The average acceleration factor for the chip $[A_{chip}(T)]$ over this time period is computed as the average of the acceleration factor for all N cores as

$$A_{chip}(T) = \frac{1}{N} \times \sum_{i=0}^{N-1} A_j(T). \quad (15)$$

The burn-in time can be estimated as the time required to achieve a given acceleration factor. Note that high NSSC implies that for a given time period T , all the cores are “on” for a high total time. In other words, the cores receive high effective stress time. Hence, with ASTPM and fast migration, each core and the entire chip accelerate fast leading to reduced burn-in time as shown in Fig. 41(a) and 41(b). Further, as ASTPM allows high NSSC for high-Vt dies, the acceleration rate is higher and burn-in time is smaller than the low-Vt corners as shown in Fig. 41(a) and 41(b).

With on-chip leakage sensors, this property can be used to reduce the total burn-in chamber time for a wafer lot and a large set of packaged dies. Fig. 41(c) shows a strong correlation between the leakage and the NSSC. Hence, using the sensed leakage values, we propose to bin the wafers or dies in different leakage bins. The maximum leakage for each bin is used to determine the required burn-in time. The binning allows the dies in the high-Vt (i.e., low leakage bins) to be stressed in a smaller time. Fig. 41(d) shows the effect of leakage-binning-based

ASTPM on the burn-in time. As expected, increasing the number of leakage bins reduces the total burn-in time for a large number of dies.

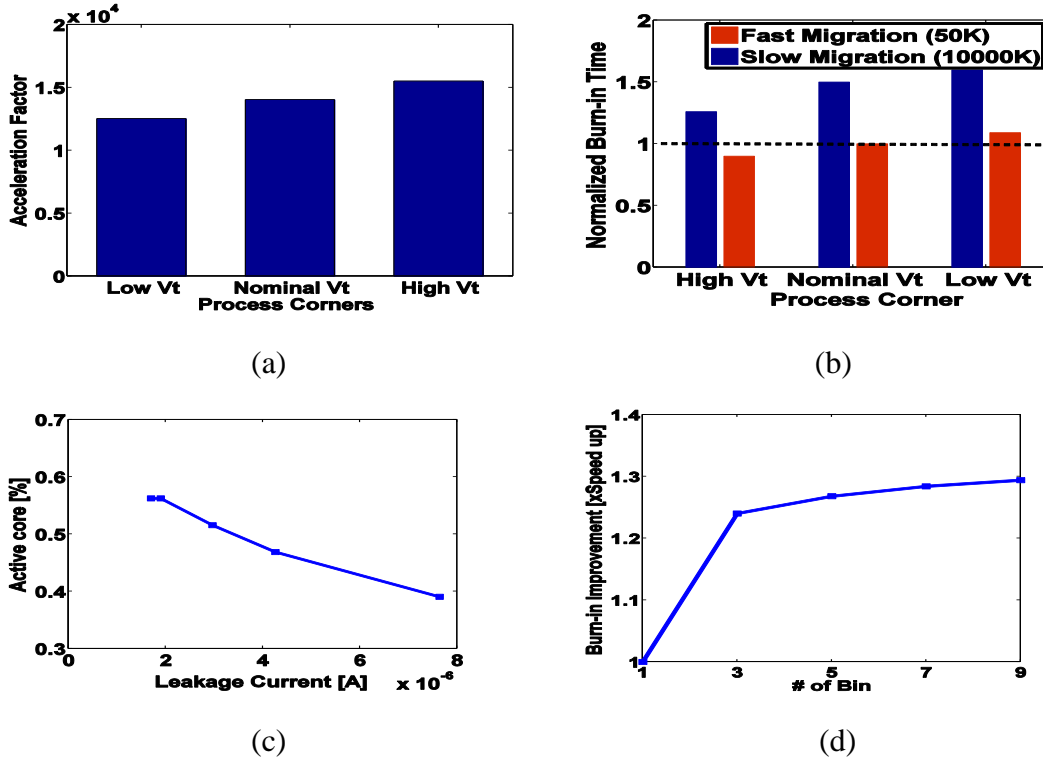


Fig. 41: Effect of ASTPM on burn-in time: (a) acceleration factor for different corners, (b) normalized burn-in time for different corners, (c) relation between NSSC and leakage current, and (d) burn-in time improvement with leakage binning.

5.7 Characterization of the Effect of Migration Interval in Normal Operation

This section presents the detailed analysis of how spatiotemporal power migration using activation and deactivation of cores can be used to minimize the peak temperature and spatiotemporal non-uniformity in normal operation of the many-core processor. In particular, this section characterizes the thermal behavior of a predictive many-core processor to show the impact of controlling migration interval (i.e., how fast migrations are performed) on the overall thermal field of a predictive 256-core processor designed in 16nm node.

5.7.1 Thermal Impact of the Migration Interval

This section studies the effect of time-slice interval considering varying computational demand

and on-core ratio (i.e., the ratio of the number of on-core to the total number of cores). The 25% to 75% variation in the on-core ratio is considered in the simulation. The random migration policy is used for the experiments. In this condition, at the end of each migration interval, the current set of active cores is deactivated, and a new random set of active core is chosen for the next time interval. The following parameters to characterize the thermal field are computed: (a) maximum chip temperature: the maximum of the temperature of all the cores at a given time, (b) spatial difference: the difference between maximum and minimum on-chip temperature at a given time, and (c) temporal difference: the difference between the maximum and minimum temperature for a core over a long time period ($\sim 10\text{ms}$) and average of that difference over all cores in the chip. For a given number of active cores, a large time-slice leads to a high maximum temperature and high spatiotemporal non-uniformity. Fig. 42 shows the thermal map at different time instances considering times-lice intervals of 100,000 (100K) and 1,000,000 (1000K) clock cycles.

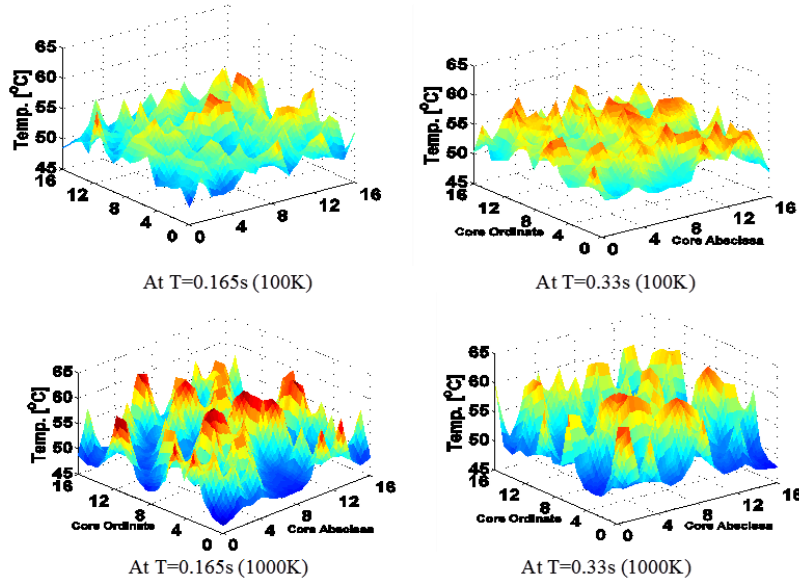


Fig. 42: Thermal behavior of many-core system.

A small time-slice shows the reduced maximum temperature as shown in Fig. 43(a). A fast

migration implies that core-to-core temperature variation is much less as shown in Fig. 43(b). The spatial difference reduces with a fast time-slice as shown in Fig. 43(c), which implies that on-chip thermal variations are more uniform with a fast migration. Both the average value and core-to-core variations of temporal difference of all cores reduce significantly with a fast time-slice as shown in Fig. 43(d). Low and similar thermal cycle for all cores implies low and similar reliability degradation for all cores.

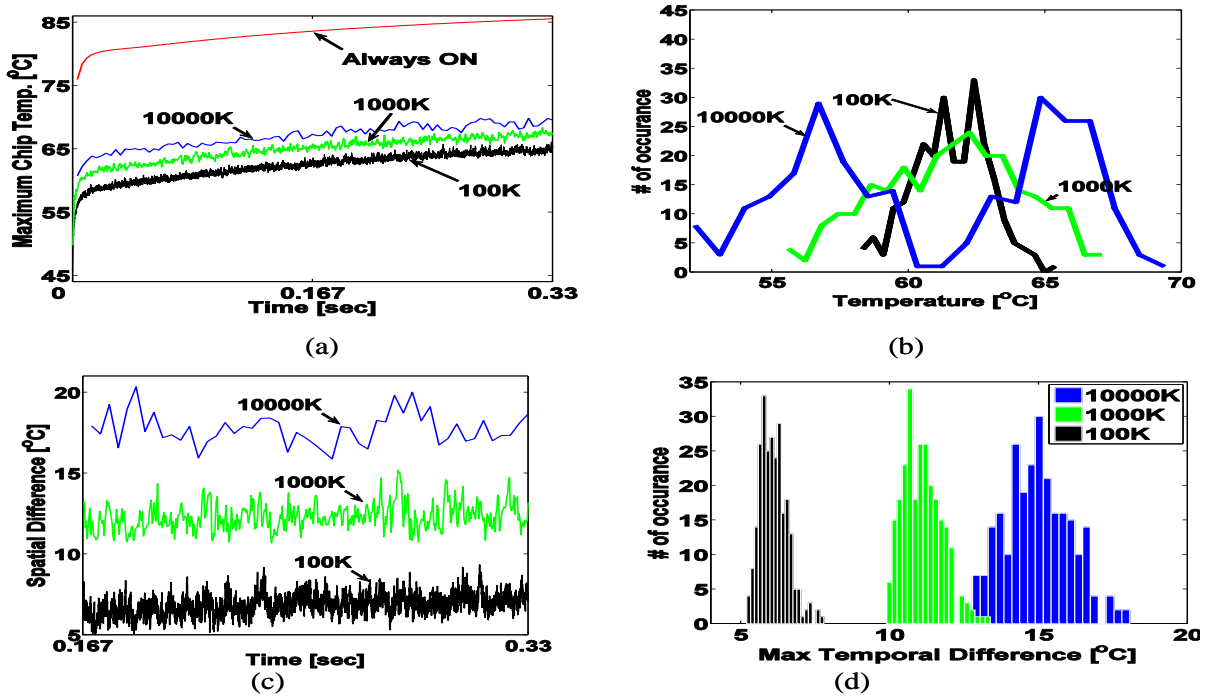


Fig. 43: Effect of time-slice on random migration (a) max temperature (b) core-to-core temperature distribution, (c) spatial difference, and (d) temporal difference. The 100K refers to 100,000 clock cycles i.e. 33 μ s for a 3GHz clock.

The reduction in the maximum temperature, spatial difference, and temporal difference with a low time-slice is effective for different numbers of active cores as shown in Fig. 44(a), 44(b), and 44(c). However, a small time-slice is required for large on-core ratio to maintain a target maximum temperature, spatial difference or temporal difference as shown in Fig. 44(d), 44(e), and 44(f). The design of the time-slice will depend on the thermal properties of the chip

particularly on the lateral resistance (for spatial redistribution) and thermal capacity (for temporal redistribution) of silicon. The time-slice can be adapted depending on the computational load or real-time temperature data.

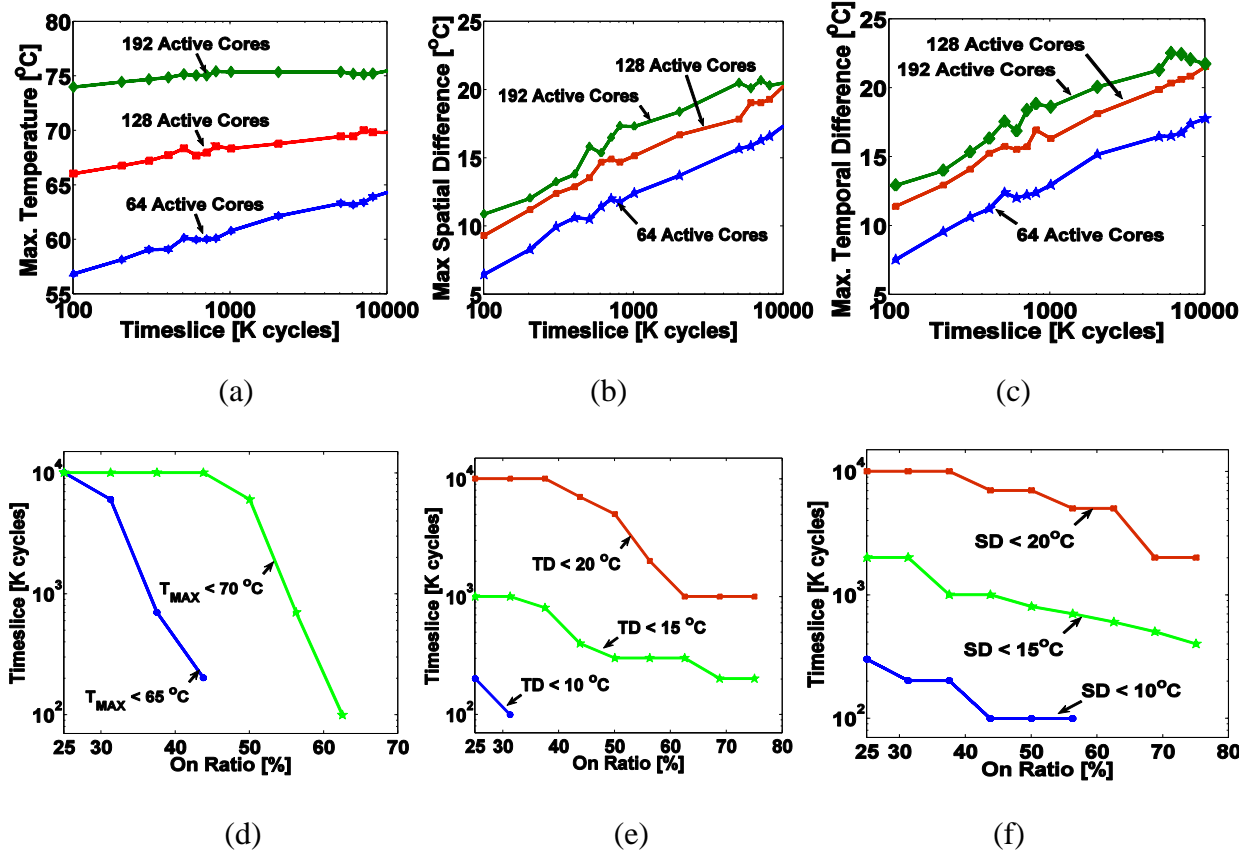


Fig. 44: Effect of time-slice on (a) max temperature, (b) max spatial difference, (c) max temporal difference, (d) time-slice for a target max temperature on ratio, (e) time-slice for a target spatial difference in different on ratio and (f) time-slice for a target temporal difference in different on ratio.

Fig. 45 shows the effect of power migration considering run-time variations in the performance demand (i.e., number of active cores). The power migration can reduce a peak temperature significantly under all demand (i.e., the number of on-cores), but the effect is more pronounced at low-to-moderate demand condition. This result suggests that power migration will help to reduce the average temperature of the chip over the lifetime reducing cooling energy as well as improving the reliability.

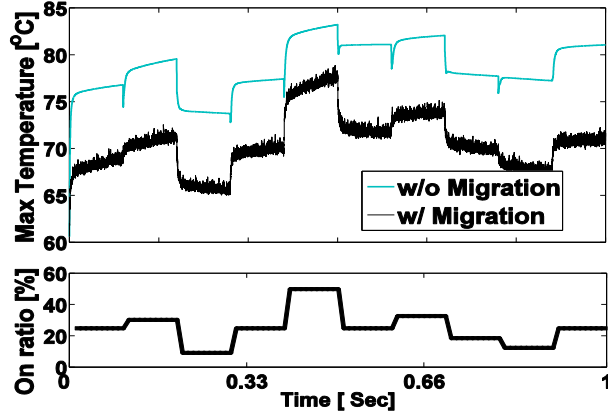


Fig. 45: Chip temperature with run-time variations in the number of active cores.

5.7.2 Electrical Impact of Migration Interval and Analysis of Tradeoffs

Although fast power migration provides a low maximum temperature and a uniform thermal map, it also requires migration of computation threads which increases the performance overhead. Therefore, an analysis of the tradeoff between thermal behavior and system performance is critical for the accurate evaluation of power migration.

5.7.2.1 Effect on Core-to-Core Delay and Leakage Variations

The delay and leakage variations are estimated considering an 8-stage FO4 ring oscillator using 16nm predictive models [31]. Fig. 46(a) and 46(b) show the histogram of delay and leakage increase compared to 45°C temperature. A low on-chip delay and leakage variation imply a better functional reliability. Fast spatiotemporal power migration helps to reduce the core-to-core delay and leakage variations. Note that at the end of a time-slice interval, the inactive cores can have lower temperature for large migration interval than the inactive cores for fast migration, as a larger migration interval provides longer time for temperature to decay. Consequently, at a high migration interval, cores have a low delay and leakage compared to the cores with fast migration. However, the operating frequency of the chip is determined by the slowest active core, and the maximum core delay is high for slow migration (i.e., high temperature of active cores). Fig.

46(c) shows the variation in the maximum delay (i.e., delay of the slowest on-core) considering on-chip thermal variation over time. Fig. 46 shows that overall chip performance for a target on-core ratio improves with fast migration. Fig. 46(d) shows the normalized leakage power for the on-cores over time. As expected, the total chip leakage reduces with a fast time-slice.

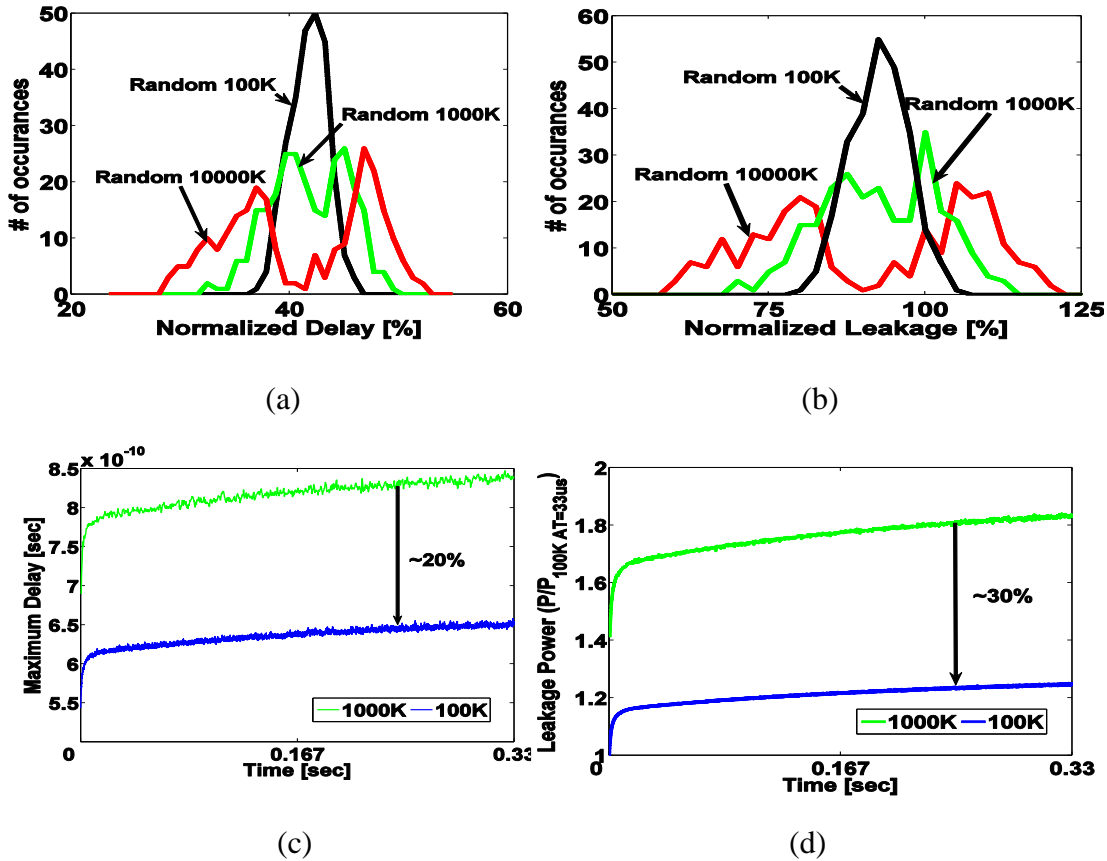


Fig. 46: Effect of power migration on (a) core-to-core delay distribution, (b) core-to-core leakage distribution, (c) chip delay (i.e., maximum core delay), and (d) chip leakage power.

5.7.2.2 The Power Overhead and Overall Power Saving

The spatiotemporal power migration will require additional transition energy for the active-inactive control. We assume during transition all internal nodes make transitions (i.e., entire core capacitance switches). The switched capacitance is estimated from the power of each core. Based on ITRS roadmap and 10% additional core-level decoupling capacitance, the transition of the supply-network capacitance is estimated. Note that for N on-cores, the migration will imply

transition overhead of $2N$ cores (i.e., N cores turn off, and N cores turn on). Additional power from migration of the flip-flop states is also considered. As the power migration is performed fast, the power overhead increases. Table 2 shows the estimated power overhead for different migration interval. The summary of the overhead analysis shows that as the power migration is performed fast, the power overhead increases. However, as the core-switching capacitance is reduced at deep nanometer nodes and migration is performed at relatively low frequency (100,000 – 1,000,000 times lower than the clock frequency), the power overhead is very low. Even at migration intervals of 100,000 clock cycles, the power overhead is limited to less than 0.1% of the total chip power. For migration intervals of 1,000,000 clock cycles, the overhead reduces to less than 0.01%. Therefore, the power overhead is not a critical challenge for power migration.

Table 2: Power analysis: power overhead of migration.

Timeslice Interval	100K	200K	500K	1000K
Power overhead	0.028	0.014	0.006	0.003

CHAPTER VI

THERMAL SENSOR DESIGN CHALLENGE

Most of thermal management methods assume accurate thermal sensors are placed in the chip. Although the thermal management methods including the ASTPM method provide better thermal field, inaccurate sensors and the lack of understanding of thermal sensor design can limit the effectiveness of the methods due to the miscalculation of temperature and performance. A temperature sensor can be divided into an analog and digital sensor. Because of low cost, a digital sensor is more widely used in recent processors. However, compared to the analog sensor, the digital sensor suffers from non-linearity and low accuracy problem. Since the thermal sensor plays an important role in the thermal management, this chapter explores design challenge of a digital temperature sensor. To study the challenge of thermal sensors, we implemented an analog BJT-based temperature sensor and a digital sensor using ring oscillators. To figure out effective logic gates as a core of the digital sensor, we have implemented ring oscillators with different logic gates. Along with the type of logic gates, we also implemented ring oscillators with different type of devices. Inverse temperature dependence issue causing difficulty in designing thermal sensor is addressed.

6.1 Test-chip Organization

The organization of a test-chip is shown in Fig. 47. Three analog BJT-based thermal sensors are placed at different location in a chip. One digital thermal sensor is located using a ring oscillator. Table 3 shows the measurement condition. The chip has been packaged using 44-pin LCC package. The total die area is $2 \times 1 \text{ mm}^2$ including the wire-bonding pads. The control block

runs off a 1.2 V supply, while V_{DD} for the test blocks were varied from 0.55V – 1.2V. The poly-resistor-based heater supply varies from 0V – 10V to heat up the chip. The maximum power density is limited to 130 W/cm².

6.1.1 Poly-resistor-based heater

To perform the temperature-related characterization of a chip, the chip needs to be placed in a thermal chamber. However, controlling the chamber temperature is not easy and increases testing cost and effort. The built-in heater reduces the testing time and cost. To mimic a maximum operating power/thermal condition of 130W/cm², a resistor of 50Ω was used as shown in Fig. 48. The power density generating temperature is controlled by varying the supply voltage across the resistor from 0 to 10V. The power density is then estimated by the measured total power dissipated by the heater and the given chip area. The heater is placed as close as possible to the test structure to reduce thermal gradient across the chip for better accuracy. Current flowing through the heater is limited to meet reliability constraints of the heater's poly and metal components [105].

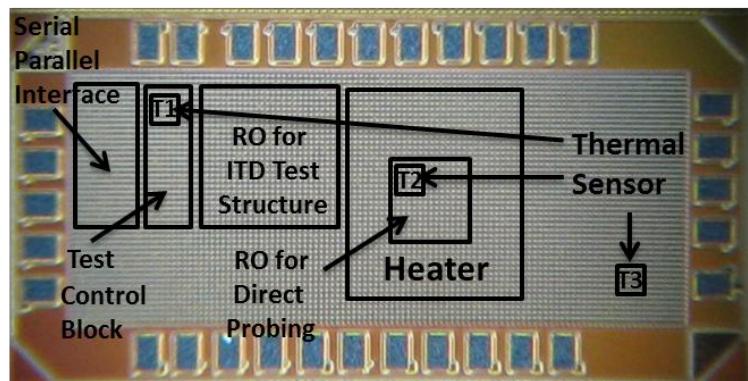


Fig. 47. The die photo showing the organization of the test chip.

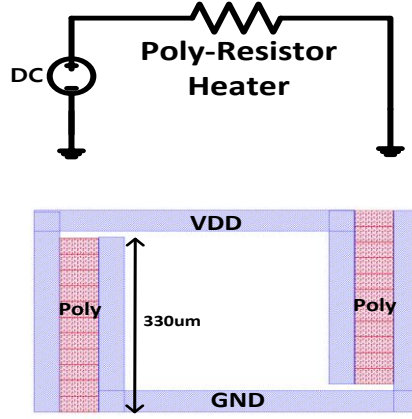


Fig. 48. Poly-resistor-based heater

Table 3. Summary of measurement conditions.

Technology	130nm CMOS
Chip size	2 mm ²
V _{CTRL}	1.2 V
V _{DD}	0.55 - 1.2 V
V _{HEATER}	0 - 10 V
Max. Power Density in Heater	130W/cm ²
Reference Clock Freq.	10KHz

6.2 Analog Sensor

To figure out the characteristics of an analog temperature sensor, a BJT-based thermal sensor was implemented. As shown in Fig. 49, the base-emitter voltage in a bipolar is measured. The base-emitter voltage shows a negative temperature coefficient, i.e., as temperature increases, the base-emitter voltage decreases. To use the negative temperature relationship, constant current source is designed and mirrored to the diode-connected BJT. Temperature of the chip is then found based on a simulated V_{BE} . Multiple temperature sensors were used in the test-chip for better accuracy.

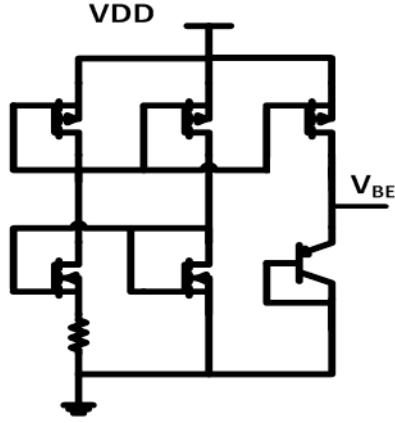


Fig. 49. Analog thermal sensor

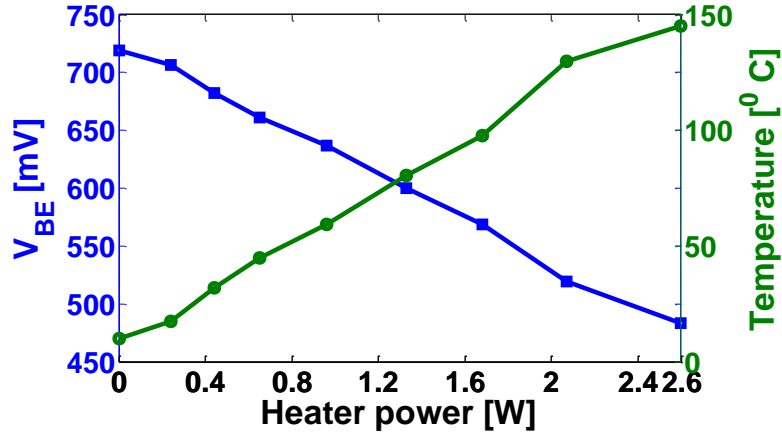


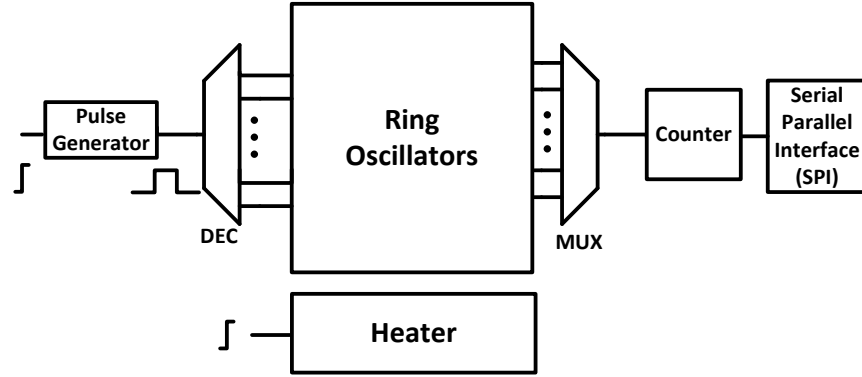
Fig. 50. Measured results showing the operation of the temperature sensor (T1) and the heater.

Fig. 50 shows the characterization of the analog thermal sensor. Measured V_{BE} decreases by $\sim 200\text{mV}$ for a heater power of $\sim 2\text{W}$ which corresponds to a simulated temperature increase of $\sim 120^{\circ}\text{C}$.

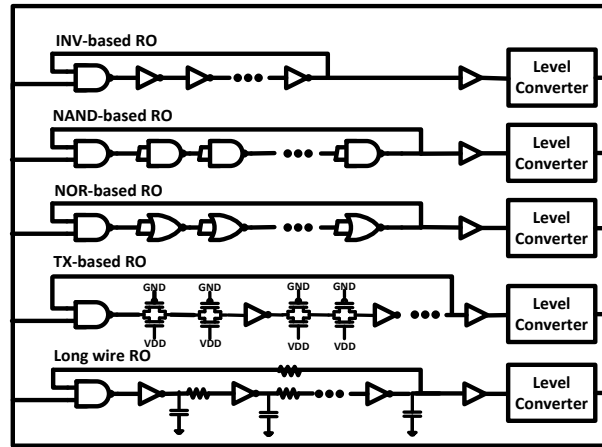
6.3 Digital Sensor

A ring-oscillator-based digital temperature sensor is designed to measure a chip temperature as shown in Fig. 51. Temperature is calculated by measuring the frequency of ring oscillators.

Ring oscillators of different gate types compose the core of the digital sensor. Control circuits include a pulse generator, a heater, level converters, and a counter.



(a)



(b)

Fig. 51. Digital sensor structure: (a) Overall block diagram and (b) detail of the ring oscillators.

6.3.1 Ring Oscillator (RO)

To assess the impact of types of gates as a core of a sensor on temperature sensing, we built multiple ring oscillators with a stage component being an INV, a NAND gate, a NOR gate, a transmission gate, and a long buffered interconnect wire. To decouple of the effect of control circuits, 101-stage ROs were used. The delay of 101-stage ring oscillator is dominant in the total delay considering the control circuits. The INV, the NAND2, the NOR2, and the NAND4 ROs

were implemented with 100-stages each plus an initial NAND2 gate. Both a nominal- V_T (NVT) and a low- V_T (LVT) INV ROs were also implemented. The gates are sized to equalize pull-up and pull-down resistances. The Transmission Gate (TX) RO was implemented using 64 TXs, 32 INVs, and a NAND2 gate such as two TXs are placed between two inverter stages. Long wire ROs was implemented with 4 INVs and a NAND2 gate, where a 400um wire (Metal 1) is fitted between two inverter stages.

6.3.2 Pulse generator and Level converter

To find the oscillation frequency of given RO, we count the transitions of the RO over a given time period. This is done as shown in Fig. 52 using a three flip-flops pulse generator that is controlled by a 10 KHz reference clock. A reference clock signal is applied to flip-flop and an inverted clock signal is applied to next two consecutive flip-flops. Once the start signal propagates through continuous flip-flops based on the clock signal, the delayed start signal by a clock signal is generated. Those delayed start signals at the output of flip-flops creates a pulse signal which has a time period of the reference clock through XNOR gate. Only test circuits are supplied with variable V_{DD} while a nominal supply is used for the control circuits. A conventional cross-coupled level converter is placed at the output of each ring oscillator, which is connected to a multiplexor.

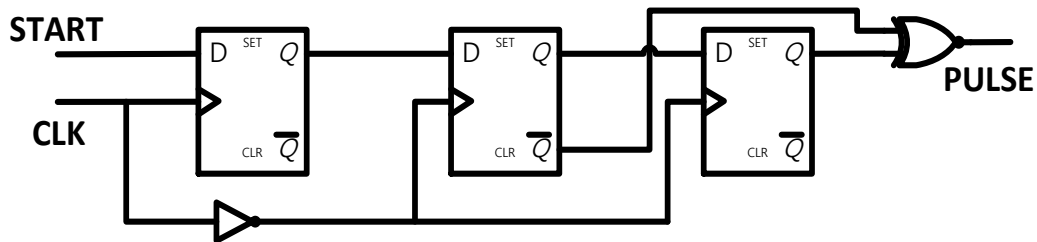
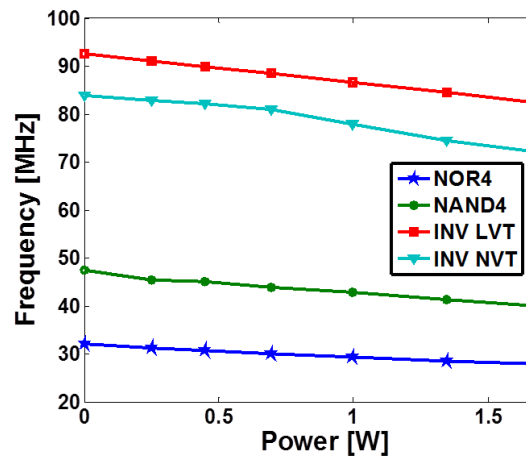
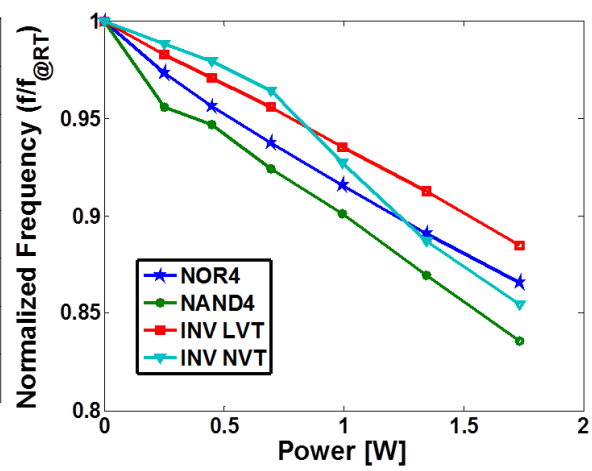


Fig. 52. Pulse generator.

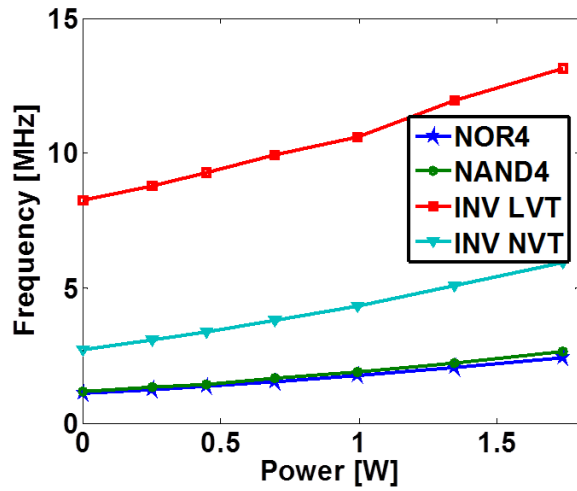
Fig. 53 (a), (b), Fig. 54 (a), (b) show the correlation between power/temperature and the frequency of four ring oscillators at nominal VDD for the digital sensor. For 130nm CMOS technology, LVT inverter-based ring oscillator show better linearity of a sensor at nominal VDD. Fig. 53 (c), (d), 54 (c), (d) show the correlation between power/temperature and the frequency of four ring oscillators at low VDD for the digital sensor. At low VDD, while all logic gates good linearity and resolution, LVT has less resolution. Change in the frequency of ring oscillators over temperature at low VDD is higher than change at high VDD. Based on these observations, nominal threshold inverter-based ring oscillator with a low supply voltage is best fit as a core of the digital sensor with respect to linearity and resolution.



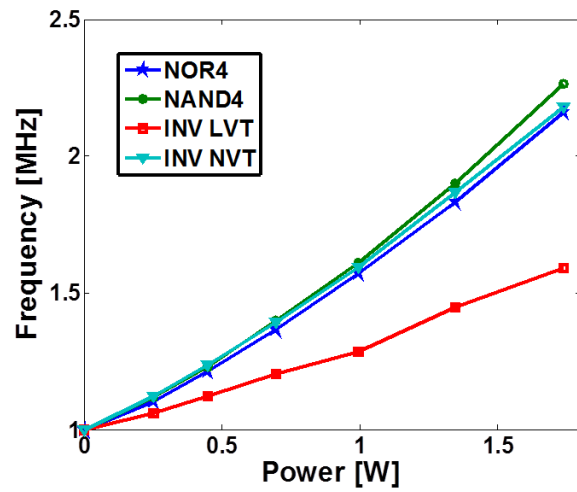
(a)



(b)



(c)



(d)

Fig. 53. Correlation between Power and Frequency of Digital Sensor: (a) correlation between power and frequency with nominal VDD, (b) correlation between power and normalized frequency with nominal VDD, (c) correlation between power and frequency with low VDD, and (d) correlation between power and normalized frequency with low VDD.

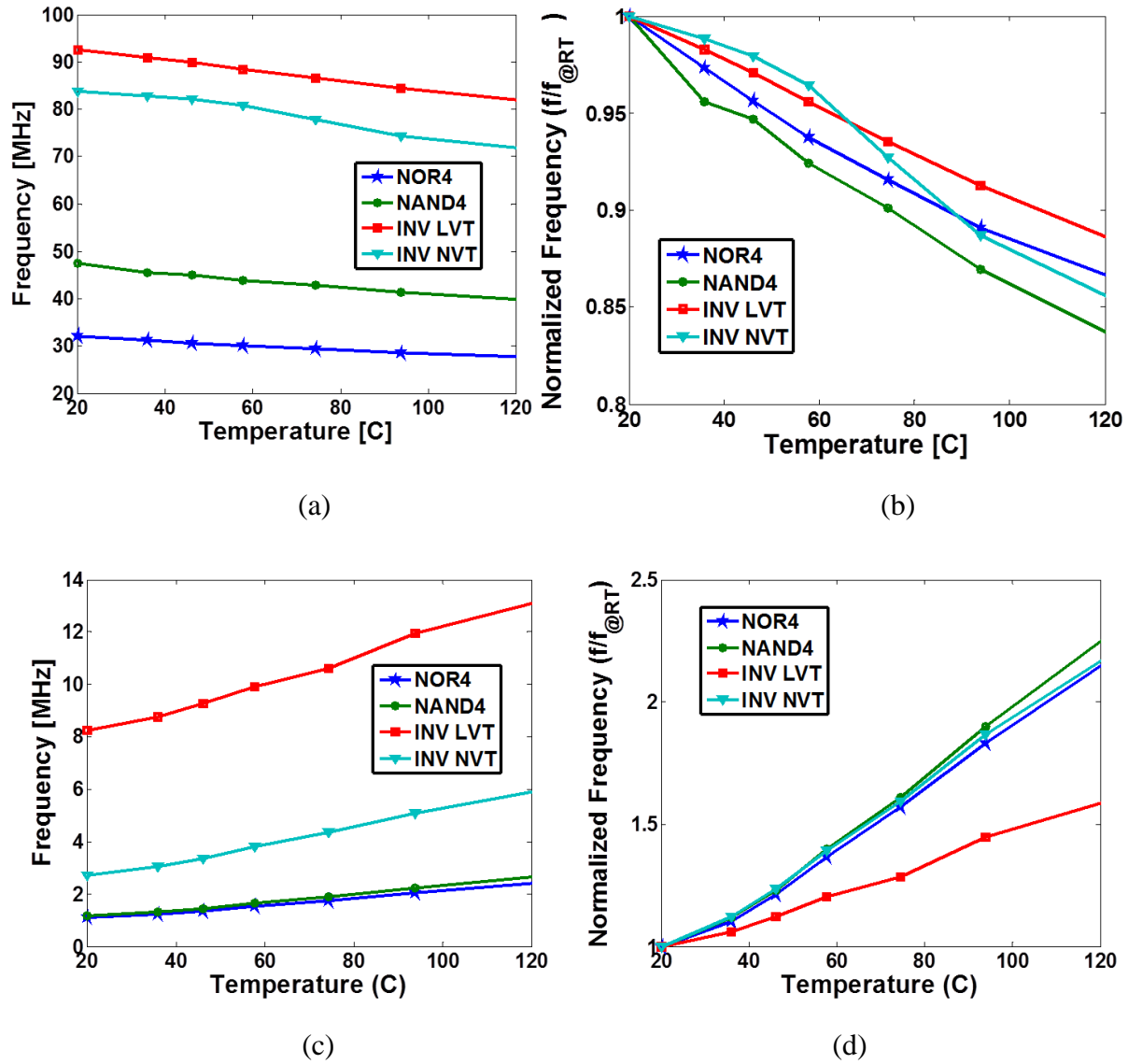


Fig. 54. Correlation between Temperature and Frequency of Digital Sensor: (a) correlation between temperature and frequency with nominal VDD, (b) correlation between temperature and normalized frequency with nominal VDD, (c) correlation between temperature and frequency with low VDD, and (d) correlation between temperature and normalized frequency with low VDD.

6.4 Design Challenge in Wide-operating Range and Low V_{DD} : Inverse Temperature Dependence

Digital circuits supporting wide V_{DD} operating range –from V_{MAX} used in high performance modes down to near threshold voltages (or NTV) used in low power modes – are emerging as

key component for next generation DVFS processors and system-on-chips. However, designing across wide V_{DD} range while ensuring correct operation under process (within-die and die-to-die) and temperature variations, is a major challenge. Temperature and voltage variations are further intertwined as temperature increase can have a positive or a negative impact on circuit delay depending on V_{DD} . At high V_{DD} , operating at higher temperature reduces device current, thus increases logic path delay. This happens as device mobility degradation is higher than V_T reduction with temperature increase. This is often referred to as the normal temperature dependence. To prevent timing errors (like setup violations) and overheating related reliability issues, clock frequency is scaled down proportionally, following a high-temperature reading from an on-die thermal sensor, for example. However, with process scaling and the introduction of high-k/metal-gate, devices exhibit higher (negative) temperature coefficient along with weaker mobility temperature sensitivity [100]. This *inverses* the impact of temperature rise on delay, particularly as V_{DD} is lowered, where a small change in V_T results in large current change. A high temperature reading from the sensor in this case can falsely indicate the need to lower frequency, while in fact the circuit delay has decreased. Alternatively, a delay based temperature sensor can incorrectly miss an overheating event if temperature dependence is not known. Understanding inverse temperature dependence thus is critical for temperature sensor design.

The delay of a CMOS digital gate driving a load C and running off a supply V_{DD} can be expressed as:

$$Delay = \frac{CV_{DD}}{I_D} \quad (16)$$

, where I_D in the linear and saturation regions is given by (17), and (18); respectively:

$$I_D = \frac{W}{L} C_{OX} \mu_n(T) \left\{ (V_{GS} - V_{th}(T)) V_{DS} - \frac{m}{2} V_{DS}^2 \right\} \quad (17)$$

$$I_D = \frac{W}{L} C_{OX} \mu_n(T) \frac{(V_{GS} - V_{th}(T))^\alpha}{m} \quad (18)$$

, where μ_n is the effective channel mobility, C_{OX} is the gate oxide capacitance per unit area, W and L are the channel width and length; respectively, and m is the body effect coefficient. The temperature dependence of mobility and V_{th} [102, 103] is respectively expressed by (19) and (20):

$$\mu(T) = \mu(T_0) \left(\frac{T}{T_0} \right)^{-\theta} \quad (19)$$

$$V_{th}(T) = V_{th}(T_0) - k(T - T_0) \quad (20)$$

, where T_0 is the initial room absolute temperature (300K), θ is the mobility temperature exponent, k is the temperature coefficient of the threshold voltage.

The elevated temperature reduces both mobility and threshold voltage. However, at higher V_{DD} , the change in V_{th} has relatively less impact on the gate over-drive ($V_{GS} - V_{th}$). Hence, the change in mobility is dominant factor resulting in a decrease in current and increase in delay at elevated temperature. However, at low V_{DD} , changes in V_{th} has a much stronger impact on the gate overdrive ($V_{GS} - V_{th}$), resulting in current increases and delay reduction at elevated temperature. The ZTC V_G point lies between the two regions and can be calculated by solving for $\frac{\partial I_D}{\partial T} = 0$. Following the method in [104], the analytical model for ZTC point over a temperature range T_0 to T_H can be expressed as follows (see Appendix A for details):

In the linear region:

$$V_{ZTC} = V_{th}(T_0) - k(T_H - T_0) + \frac{m}{2}V_{DS} + k\left(\frac{T_H}{\theta}\right) \quad (21)$$

In the saturation region:

$$V_{ZTC} = V_{th}(T_0) - k(T_H - T_0) + \alpha k\left(\frac{T_H}{\theta}\right) \quad (22)$$

This model shows that ZTC point is different for linear and saturation regions and is dependent on process parameters and operating condition. For the process example of $k=2.4\text{mV/K}$, $m=1$, $\alpha=2$, and $\theta=1.5$, $T_H = 400\text{K}$ [103], the ZTC point in saturation region is higher than in the linear region. The linear ZTC gets further smaller with lowered V_{DS} as given in Equation (21).

6.4.1 Measurement Results

The test chip has been fabricated in 130nm CMOS. Fig. 47 shows the die-photo and the organization of the test-chip. The test-structure shown in Fig. 51 was implemented. A Serial-to-Parallel Interface (SPI) block was used to control the measurement. The SPI block selects the ring-oscillator to test by applying the signal to a decoder in the test-chip. The SPI block also collects a transition counter number, which can translate into the frequency of the ring-oscillator, from the counter in the test-chip. The SPI software [107] helps to control the SPI block from a computer. Along with the test-structure shown in Fig. 51, ring-oscillators were also included for direct probing using an oscilloscope.

For measuring the ZTC point, the heater power was varied and the V_{DD} at which the measured frequency is independent of the heater power was measured as the ZTC point. The temperature sensors were used only for reference and do not impact the ZTC measurement accuracy.

Fig. 55 shows the waveform of the inverter-chain ring oscillator at different V_{DD} measured at high and low (room) temperature using direct probing. It has clearly shown that for nominal voltage (1.2V for 130nm CMOS technology), the frequency decreases with temperature rise, while at 0.55V which is below the ZTC point, the frequency increases with temperature. Fig. 56 summarizes the temperature dependence of the measured frequency of the inverter chain based ring oscillator for different V_{DD} . We observe a ZTC point of approximately 0.8V.

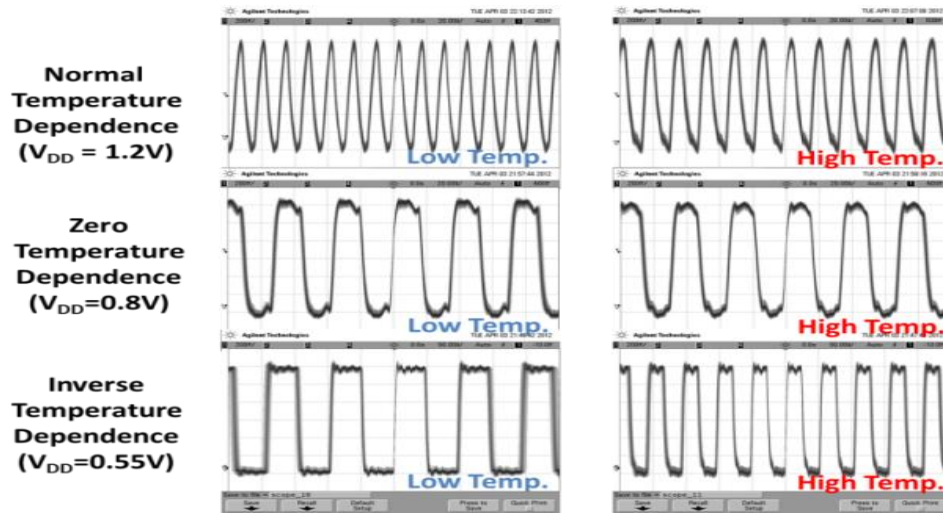


Fig. 55. Inverter chain RO waveform at different V_{DD} and temperature.

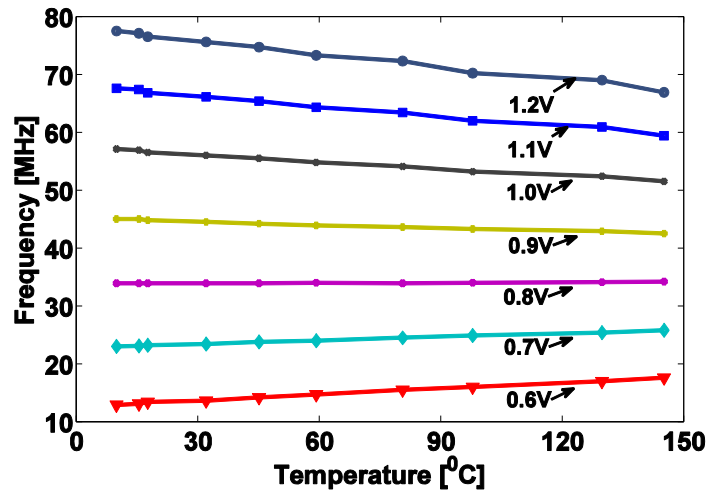


Fig. 56. Measured results showing temperature dependence of an inverter chain RO.

6.4.2 The impact of device type

Two INV ring oscillators were implemented: one with nominal-threshold (NVT) devices and the other with low-threshold devices (LVT) to estimate the impact of device type on ZTC. As expected, the LVT RO frequency is higher than that for the NVT RO at ISO V_{DD} , as shown in Fig. 57. Since the ZTC point is proportional to V_T at room temperature (Equation. 20), LVT INV RO has a lower ZTC point.

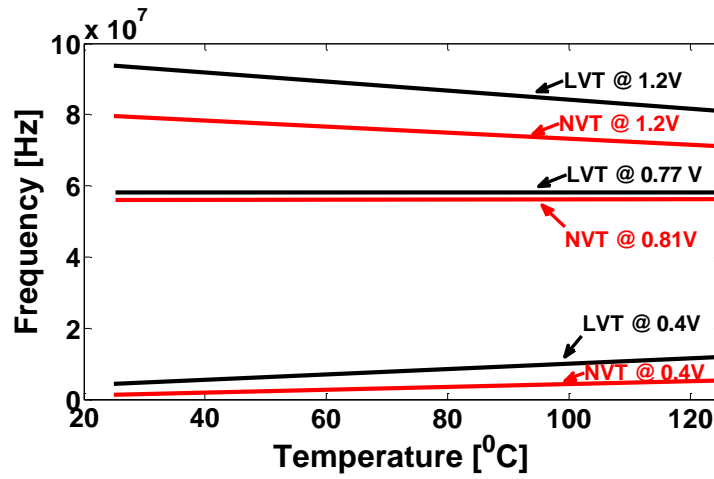


Fig. 57. Impact of the LVT and HVT devices on ZTC.

6.4.3 The impact of circuit type

Table 4 shows the average ZTC point for the different RO's measured over a number of dies.

Table. 4. Impact of type of devices on temperature dependence.

Type of circuits	Average ZTC point [V]
NVT Inverter	0.783
LVT Inverter	0.770
NAND2 (NVT)	0.767
NAND4 (NVT)	0.736
NOR2 (NVT)	0.789
Transmission gate (NVT)	0.819
Long wire (NVT)	0.866

The difference in ZTC points is explained as follows. According to Eq. (21), (22), the ZTC point varies from linear to saturation region. For the NAND-2 RO, a portion of the NMOS pull-down network operates more in the linear region during switching, thus its ZTC point is dominated by the linear region ZTC. As the stack height increases ($\text{NAND4} > \text{NAND2}$), V_{DS} across the devices further reduces and a larger portion of the gates operate in the linear region making the ZTC point of the NAND4 RO lower than that of the NAND2 RO. In this technology, since the room temperature threshold voltage of NMOS is lower than that of PMOS ($V_{TN} = 0.37\text{V}$, $|V_{TP}| = 0.4\text{V}$), a 2-NAND RO has a lower ZTC point than the 2-NOR RO. The transmission gate (TX) RO has a higher ZTC point than the inverter RO.

The long wire RO includes wire capacitance between gates. During transition, larger output capacitance causes transistors of the driver gate to operate longer in saturation. The higher capacitance also increases the slew. Hence, during signal transition, V_{GS} of the devices in the receiving gate remains lower resulting in a higher sensitivity of delay to any V_{th} change. Thus, long wire has higher ZTC points for this process. The length of the wire segment in this measurement was limited to $\sim 400\mu\text{m}$ resulting in a small wire resistance. Therefore, the ZTC is mainly determined by the temperature dependence of devices. One nice insight here is that an interconnect-dominant path (e.g. a bus or a clock network) will have different temperature dependence than a gate dominant path. This needs to be considered during design to reduce overall temperature margin.

6.4.4 The Impact of process variation

To evaluate the impact of process variation on the ZTC point, multiple (20) dies were measured. The chip-to-chip spread in the ZTC points are shown in Fig. 58. The simulated values

of the ZTC points at typical corner are also included in the figure for illustration. The correlation between chip-to-chip variations the measured operating frequency of the ROs and the ZTC points are also studied. The above correlation for the inverter chain RO is shown in Fig. 59. In general, a good correlation is observed - chips to higher threshold voltage (lower frequency) have higher ZTC point. The spread in the ZTC point for each circuit path is up to ~100mV. But the spreads for different circuit paths are different. For example, the long wire RO, the chip-to-chip variation in ZTC is more possibly due to the fact that less WID variation averaging happens as this RO has only five gates compared to other ROs. When both circuit type and process variation are considered, the difference between the minimum ZTC (NAND4) and the maximum ZTC (long wire) value of ZTC points is ~200mV.

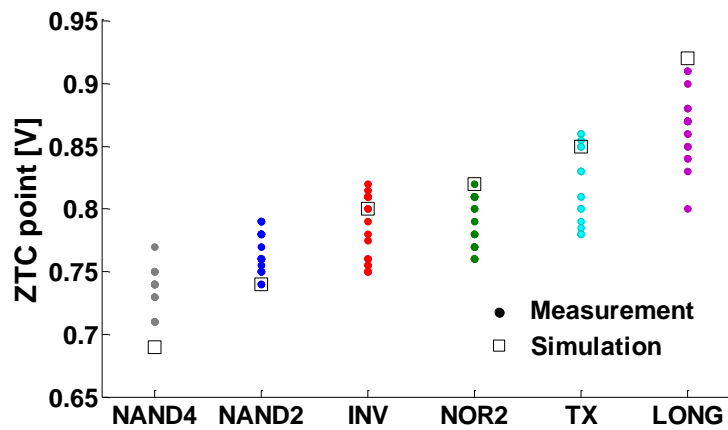


Fig. 58. Measured ZTC points for different circuit paths considering process variations. The data from 20 chips are shown.

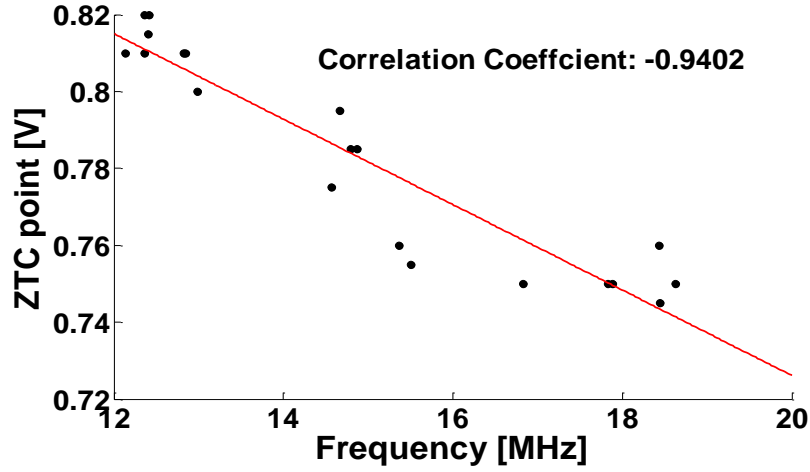


Fig. 59. Measured correlation between ZTC points and frequency for the inverter chain based ROs.

6.5 Conclusion

This chapter verified the characteristics of analog and digital thermal sensors through a test-chip. For a digital sensor, the impact of supply voltage and types of ring-oscillators on linearity and resolution is analyzed. Based on the measurement from the test-chip, nominal threshold inverter-based ring oscillator with a low supply voltage is best candidate as a core of digital sensor with respect to linearity and resolution. Inverse temperature dependence issue causing difficulty in designing thermal sensor is addressed. The impact of device, type of gates, and process variation on temperature dependence is observed.

CHAPTER VII

POST-SILICON THERMAL SYSTEM IDENTIFICATION AND PREDICTION

In Chapter V, we discussed the adaptive spatiotemporal power migration (ASTPM), which improves the thermal field. The existing methods to compensate for variations in temperature over space and time including the ASTPM are designed and the effectiveness of those methods are verified based on the pre-characterized thermal systems. However, the characteristics of thermal systems vary due to the manufacturing imperfection, which can cause variations in the effectiveness of methods. Therefore, the thermal system identification and temperature prediction considering electrical and thermal variations are necessary for efficient thermal management.

7.1 Introduction

Characterization of the spatiotemporal variation of the on-chip junction temperature (*the transient thermal field*) is crucial for thermal-aware design, assembly, and management for reliable in-field operation of a chip (die and package) [71, 72]. The thermal field is generated by the interaction of time-varying power pattern and the thermal properties (resistivity and heat capacity) of die and package materials. Further, the thermal properties of the die/package assembly [e.g. conductivity of thermal interface materials (TIM)] can vary between different instances of same IC (chip-to-chip variation) or over time (e.g. delamination in TIM [73]). Moreover, imperfections in the manufacturing process leads to die-to-die and within-die process variations in transistor leakage [71]. The leakage and temperature are positively correlated – a higher temperature results in higher leakage which further increase the temperature. Hence, for

same dynamic power, chip-to-chip leakage variation leads to variation in on-chip temperature [74].

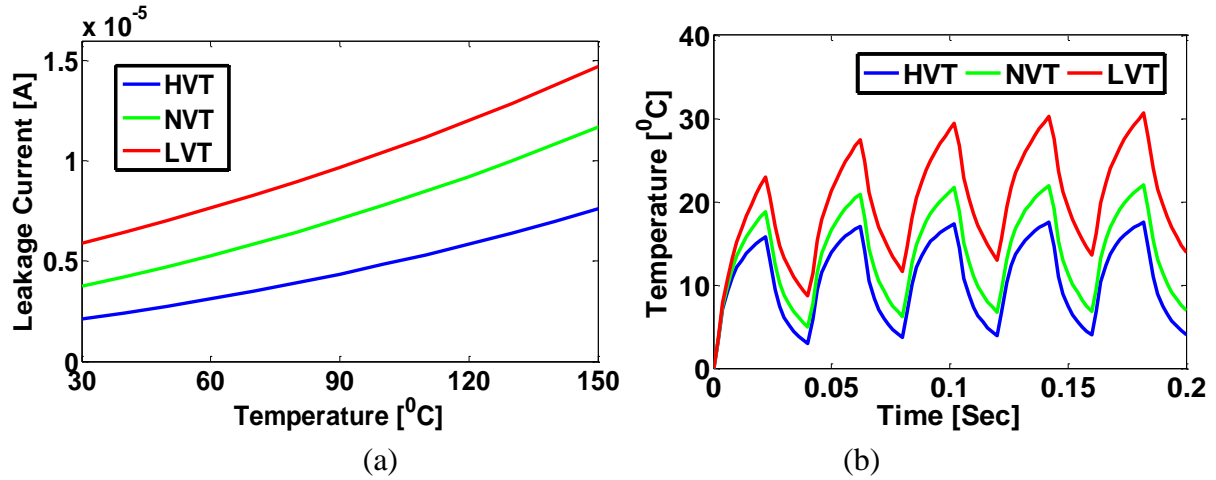


Fig. 60: Illustration of the need for post-silicon transient thermal analysis considering process variation: (a) the interaction of an average for all input condition and temperature in a NAND2 gate considering different process corners (HVT – high threshold voltage, NVT – nominal threshold voltage, and LVT – low threshold voltage corner). (b) the effect of such interaction for an example self-consistent thermal simulation (using distributed RC network) considering a square wave dynamic power profile (e.g. turning on and off a chip after a time-interval) and leakage of 10 million NAND2 gates.

Fig. 60 illustrates the impact of process variation and leakage-temperature interaction on thermal behavior of a chip using example simulations in predictive 22nm node. As the die-to-die process variation increases with technology scaling, the post-silicon chip-to-chip variation in transient thermal field is also expected to increase. This challenge is further enhanced by many-core processor architectures running increasingly data intensive and unstructured workloads. As the power, performance, and lifetime reliability of processors depend on the transient temperature, in-field reliable operation of many-core processors needs the accurate characterization of the interaction of workload variation and chip-to-chip/package-to-package variations in thermal/electrical properties. This leads to a new challenge - **post-silicon prediction of the transient thermal field**. The objective of *post-silicon thermal prediction* is to predict the transient temperature of a particular instance of a packaged IC for various workload and

considering chip-to-chip and package-to-package variations in electrical (leakage) and thermal properties.

7.2 Contributions

This section presents a unique approach for transient thermal analysis that addresses the specific requirements of post-silicon thermal prediction. The proposed approach, referred to as Thermal System Identification or TSI, is based on principles of system identification, frequency domain signal analysis, and positive feedback system. We develop the mathematical principles of the proposed approach and demonstrate its effectiveness in post-silicon thermal analysis of a 64-core processor at predictive 22nm node [83]. Each core of the processors is modeled as close to the Intel Nehalem [84] architecture running at 3.0GHz. The post-silicon characterization of a multicore chips can be used by operating systems to schedule workloads since the identification of the chip thermal system enables schedulers to reason about the thermal consequences of scheduling a specific workload on a target chip. This understanding can also be exploited in configuring large system (e.g., data centers) via thermally compatible aggregations of multicore packages. Fig. 61 shows the overall flow of the proposed post-silicon thermal prediction approach. We first extract the system transfer function using the controlled power/thermal measurement which will be discussed in Section 7.4 and Section 7.5. For the power profile to predict a temperature, we compute the fast Fourier transform (FFT) to obtain the power spectra $P(w)$. From the power spectra, we next compute the frequency response of the temperature by multiplying the power spectra and the extracted system transfer function spectra. The inverse Fourier Transform of the temperature responses is used to compute the time-domain variation of temperature at each location.

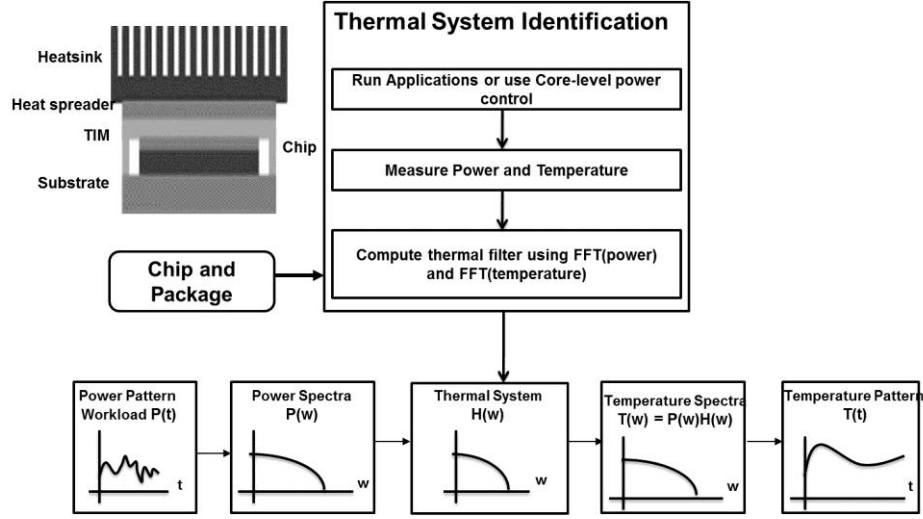


Fig. 61: Overall methodology of post-silicon prediction of the transient thermal field. The method uses the time-frequency duality to extract thermal system in frequency domain using post-silicon measurement and use that to predict transient temperature profile.

This chapter makes the following contributions:

- **High-level Transfer Function of the Thermal System including Leakage-Temperature**

Interaction: We provide a high-level abstraction of the thermal behavior of a chip as a multi-input multi-output (MIMO) system where power sources are system inputs and observed temperature values at different locations are the system outputs. The interaction of leakage and temperature is used as an integral part of the high-level MIMO system. We show that the thermal system can be represented in frequency domain as a filter matrix. In time domain heat diffusion equation represents a distributed-RC network which behaves as a low-pass filter in frequency domain. This is augmented with a positive feedback path representing leakage-temperature interaction.

- **Thermal System Identification - Post-silicon Extraction of Transfer Function of the Thermal System and Fast Prediction of Transient Thermal Field:**

We present methodologies that can identify the thermal system (i.e., the thermal filters) after fabrication and packaging using

sequences of on-chip power and temperature measurements. These methods allow one to construct a unique thermal system for each chip (thermal system identification or TSI). We present methods to accurately predict the chip-specific transient thermal fields for varying workloads using the corresponding thermal filter matrix $[\mathbf{H}(\omega)]$. The frequency response of the temperature variation over a time interval is computed from the Fourier transform of power pattern in that interval and the filter matrix $[\mathbf{T}(\omega)=\mathbf{H}(\omega)\times\mathbf{P}(\omega)]$. The time-domain temperature is obtained from the temperature spectra.

• **Hardware Validation – Extraction of Transfer Function of the Thermal System for a Test-chip Temperature and Fast Prediction:** We build a test-chip to extract the transfer function of the thermal system and predict transient temperature of the chip. Extraction of the thermal filter of a test-chip is performed by applying power signals of different frequencies and measuring the temperature from in-built temperature sensors. The extracted filters were used to predict temperature for arbitrary power patterns in the on-chip heater and verified against measured temperature.

7.3 Related work and Novelty

The existing transient thermal simulation methods (finite element/volume or distributed RC), suitable for fine-grain design time transient thermal analysis, require accurate estimation of thermal resistivity and heat capacity of all materials [75-77]. Many papers have studied on how to measure the thermal resistance and capacitance of thermal interface material (TIM), heat sink, convective, and heat spreader [78-81]. Many steady-state methods are modeled after ASTM D5470 [78]. A. Poppe et al. presented dynamic electrical temperature measurement [79] and R. Campbell et al. presented the flash diffusivity method for accurate measurement of thermophysical property data [80]. The measurements of thermal resistance and capacitance

suffer from repeatability, contamination, pressure, and inaccuracy problems. Even if we measure accurately the thermal resistances of TIM, heat sinks, and interface, in stacking condition, those values are changed due to imperfect attachment and manufacturing. K. Kurabayashi et al. presents that the die attach resistance differs substantially from the value predicted using the bulk thermal conductivity of the attachment material because of partial voiding and delamination [82]. Consequently, the fine-grain distributed-RC-based thermal simulators used during design time are difficult to adopt for post-silicon thermal analysis.

Several methods have been proposed in recent years for fast steady-state spatial thermal map (e.g., power blurring method in [85] and discrete cosine transform (DCT)-based method in [86]), fast transient temperature simulations [87-88], and fast spatiotemporal analysis considering multilayers of power and materials (e.g., ThermalScope [89]). The TSI-based approach provides important advantages in post-silicon thermal analysis over the above mentioned approaches used in fine-grain design-time thermal analysis. First, the proposed approach performs temperature prediction using the thermal transfer function extracted from the full thermal system (i.e., stacks of heat sink, spreader, TIM, and chip), instead of computing thermal resistance and capacitances of individual materials in isolation. Therefore, the effects of any non-uniformity and/or uncertainty in the thermal properties of the materials are captured in the extracted transfer function. Moreover, as the leakage temperature interaction is considered as a part of the MIMO system, the effect of process variation of individual chips is also automatically considered. Second, the fast simulators mentioned earlier do not consider leakage-temperature interaction. Currently, the transient temperature estimation considering leakage-temperature interaction is performed using distributed- RC-based simulators (e.g., Hotspot [91]) where leakage power is updated in each time-step based on the current thermal map [89-92]. Therefore, higher accuracy

of the temperature estimation requires fine-grain time-step which in turn increases simulation time. In the proposed approach the leakage-temperature interaction is incorporated in the system transfer function and temperature estimation is performed in the frequency domain. Consequently, the accuracy of the proposed method is less sensitive to time-step allowing fast estimation of transient temperature.

7.4 Mathematical Approach

In this section, the mathematical principles of the proposed thermal modeling are explored. In thermal RC-based model, the temperature at a location is the analog of voltage across the thermal capacitance at that location, and a current source represents the power dissipation causing heat generation. Considering the thermal system as a black box, we can simply consider a power and a temperature as an input and an output of thermal system, respectively.

7.4.1 Modeling the MIMO Thermal System with Leakage-temperature Iteration

From a single core or a chip with only one deterministic hotspot to many/multi core, the number of input power sources and temperature observation points increases. In a many-core system, the temperature of a core is affected by the multiple input sources (i.e., power profile of multiple cores). We need to consider the transfer functions from multiple source cores to observation core. We thus can consider many-core system as a multiple-input multiple-output (MIMO) system where the temperature of an observation point is affected by the multiple input power sources. Since a distributed-RC network is a linear system, superposition principle can be applied, i.e., the temperature at one location is the additive response of all power sources in the system. Assume that there are M power sources organized into $m \times m$ 2D grids. We further assume that there are L numbers of observation points organized in $l \times l$ grids. The temperature at the observation point (i, j) in frequency domain can be estimated as:

$$T_{ij}(k) = P_{11}(k)H_{11 \rightarrow ij}(k) + P_{12}(k)H_{12 \rightarrow ij}(k) + \dots + P_{mm}(k)H_{mm \rightarrow ij}(k) = \sum_{p,q=1}^m P_{pq}(k)H_{pq \rightarrow ij}(k) \quad (23)$$

Note $H_{ij \rightarrow ij}(\omega)$ is defined as the self-transfer function of a location (i.e. the transfer function connecting power and temperature of a location (H_{self})). Likewise $H_{pq \rightarrow ij}(\omega)$ ($\forall p, q \neq i, j$) is defined as the cross transfer function (H_{cross}) that connects power of one location and temperature of another. The above formulation leads to the 2D filter matrix for the MIMO system (Fig. 62):

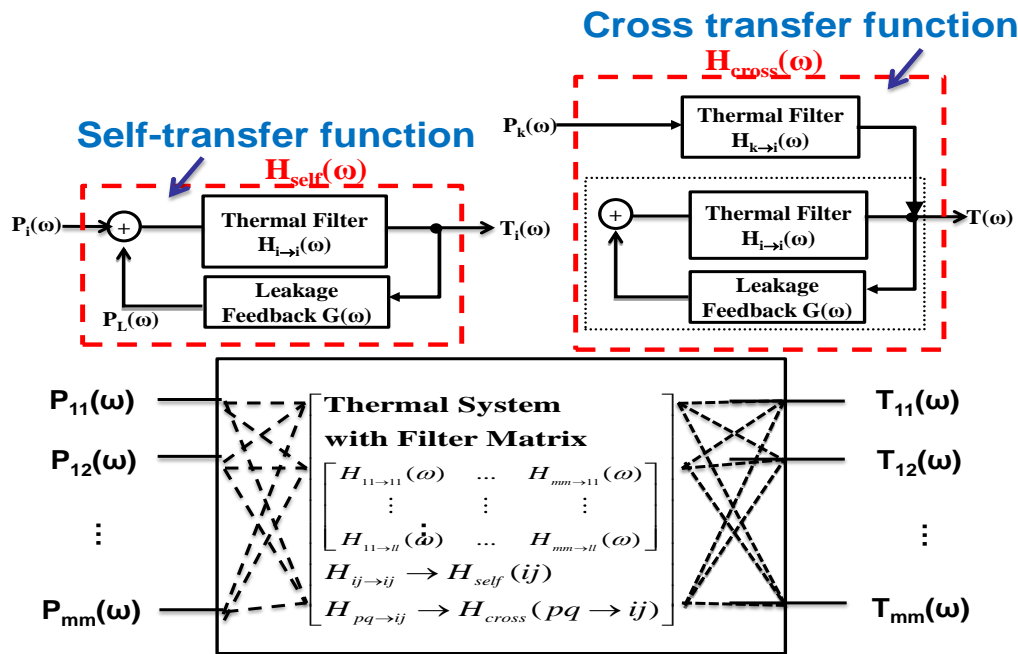


Fig. 62: Mathematical principle of the proposed approach. The thermal system is considered as a MIMO system.

$$\begin{pmatrix} T_{11}(\omega) \\ \vdots \\ T_{ll}(\omega) \end{pmatrix} = \begin{bmatrix} H_{11 \rightarrow 11}(\omega) & \dots & H_{mm \rightarrow 11}(\omega) \\ \vdots & \ddots & \vdots \\ H_{11 \rightarrow ll}(\omega) & \dots & H_{mm \rightarrow ll}(\omega) \end{bmatrix} \begin{pmatrix} P_{11}(\omega) \\ \vdots \\ P_{mm}(\omega) \end{pmatrix} \quad (24)$$

We now estimate the self and cross transfer functions considering the leakage feedback. Without loss of generality, we explain this considering two sources and two locations. Consider leakage current (P_L) depends on temperature as:

$$P_L(T) = P_L(T_0) + f(T) \quad (25)$$

,where $P_L(T_0)$ is the leakage power at room temperature, and the function $f(T)$ represents sensitivity of leakage power to temperature. First, we consider H_{self} (i.e. the temperature of location i due to the power source of at location i). We obtain:

$$T_i(\omega) = \left[P_{D_i}(\omega) + P_{L0_i}(\omega) + F(f(T_i(t))) \right] H_{i \rightarrow i}(\omega) = \left[\underbrace{P_{D_i}(\omega) + P_{L0_i}(\omega)}_{P_i(\omega)} + \alpha T_i(\omega) \right] H_{i \rightarrow i}(\omega) \quad (26)$$

, where $P_D(\omega)$ is the frequency response of the dynamic power, $P_L(\omega)$ is the frequency response of the leakage power, α is the sensitivity coefficient, and $F(f(T(t)))$ is the frequency response of the sensitivity of leakage power to temperature. The last approximation assumes a linear interaction between leakage and temperature to improve analytical tractability (i.e., $f(T(t)) = \alpha T(t)$). Both the room temperature leakage (P_{L0}) and the coefficient (α) depend on leakage-temperature interaction. Note $P_i(\omega) = P_D(\omega) + P_{L0}(\omega)$ is the spectral response of power without leakage-temperature feedback (can be estimated from the workload). Now the thermal system model can be represented as (Fig. 62):

$$T_i(\omega) = \left[\underbrace{\frac{H_{i \rightarrow i}(\omega)}{1 - \alpha H_{i \rightarrow i}(\omega)}}_{H_{\text{self}}} \right] P_i(\omega) \quad (27)$$

We now evaluate the temperature of location i due to power source at location k . We apply superposition principle during this evaluation estimate $T_i(\omega)$ assume $P_i(\omega) = 0$. However, the heat generated in location k propagates to location i which increases the temperature of location i . Increase in temperature at location i triggers the leakage feedback loop at location i . This results in leakage power at location i and hence, increases temperature of location i . The temperature increase in core i due to power of core k is therefore estimated as (Fig. 62):

$$T_i(\omega) = P_k(\omega)H_{k \rightarrow i}(\omega) + \underbrace{\alpha T_i(\omega)H_{self}(\omega)}_{H_{cross}} = \left[\frac{H_{k \rightarrow i}(\omega)}{1 - \alpha H_{self}(\omega)} \right] P_k(\omega) \quad (28)$$

7.4.2 Methods for Thermal System Identification

The principle discussed above requires frequency response of the self and cross transfer functions for each chip (i.e. TSI). To perform TSI on the MIMO system, one input power source is excited at a time, and temperature is measured at all observation points considered. Hence, the equation (24) transforms to:

$$\forall i \text{ \& } j: T_{ij}(\omega) = P_{pq}(\omega)H_{pq \rightarrow ij}(\omega) \Rightarrow H_{pq \rightarrow ij}(\omega) = T_{ij}(\omega) / P_{pq}(\omega) \quad (29)$$

The above equation can be used to estimate the thermal filter from all inputs power sources to all temperature observation points. As equation (29) is division of two complex numbers, both magnitude and phase of the filter response are extracted. For better accuracy, it will be efficient to minimize the leakage of unselected locations.

7.5 Applications of TSI to Thermal Modeling of Many-Core Processors

In this section, we apply the TSI-based approach to the post-silicon thermal prediction of many-core processor. We consider one temperature sensor is present in each core. Therefore, the MIMO thermal system for many-core processors has power of each core as an input and temperature of each core as an output.

7.5.1 Baseline Thermal Simulator used for Verification of the Proposed Approach

We first describe the baseline thermal simulation platform used to verify accuracy of the TSI based approach. We consider 3D model of the thermal system including chip, TIM, heat spreader, and heat sink. 3D distributed RC grid is generated for the different regions of the system. We use

circuit simulator, HSPICE, for solving the distributed RC grid in time domain. The power profiles are applied as current sources. The chip is modeled as a homogenous 64 core processor with private cache designed in predictive 22nm technology (total chip area 400mm^2 , each core and private cache $\sim 6.25\text{mm}^2$). Each core was modeled as close to Intel Nehalem architecture [84] running at 3.0GHz. Power traces of SPEC 2006 benchmark suites are generated using cycle-accurate architecture simulation for timing (Zesto [93]) and power (McPAT [94]) considering x86 architecture [108]. Each benchmark was run or repeated for 0.5 seconds in real time. The above environment considers architectural inputs (e.g, cache sizes, instruction decode width, number of execution units, etc.) and device parameters at various technology nodes to estimate the physical features of the processor. The example power traces obtained from the simulation are shown in Fig. 63 (a).

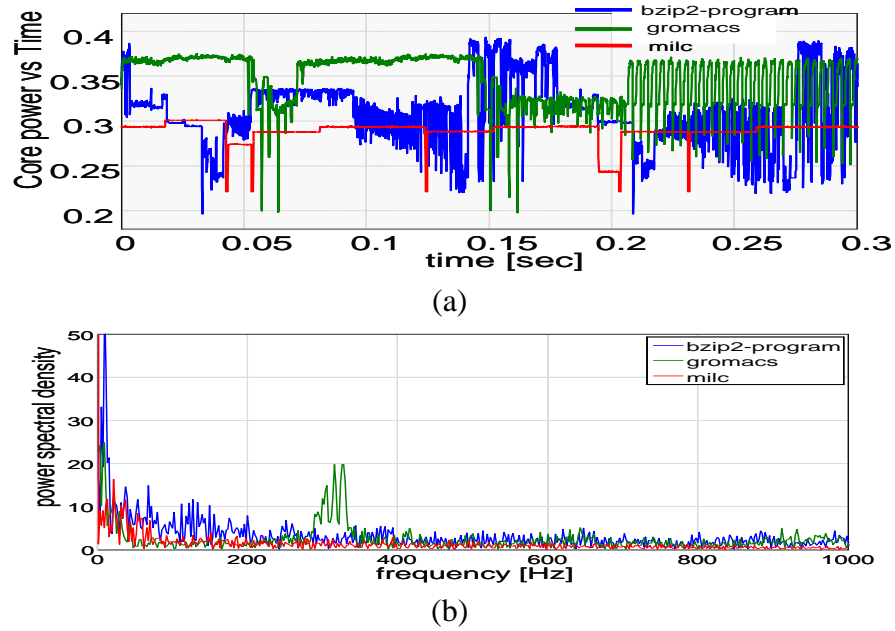


Fig. 63: (a) Transient power traces of exemplary benchmarks for SPEC2006 applications (b) The frequency response of the power traces.

7.5.2 Thermal System Identification for Many-core processors

The practical challenge in TSI of many-core processors is the generation of power spectra in

equation (29). The accurate approach is to apply sinusoidal power waveforms of different frequency (small signal analysis). However, generating sinusoidal power waveform in hardware (in a chip) is challenging. We propose two alternative approaches.

Power Spectra Generation with Core-Gating Control: First, we propose to control the core-level power and clock gating (i.e. core-gating available in current processors [95-96]) to generate power patterns of desired frequency spectra. To illustrate this approach, we perform SPICE simulation considering core gating (Fig. 64). We consider the core as hundreds of 15-stage ring oscillators to emulate dynamic power. Each core is controlled with a periodic sleep control signal of a given frequency which generates periodic power pattern of same frequency. Hence, by controlling the period of the sleep control signal we can modulate the spectral behavior of the generated power patterns. The on-chip power monitors can be used to sense the core-level power [97].

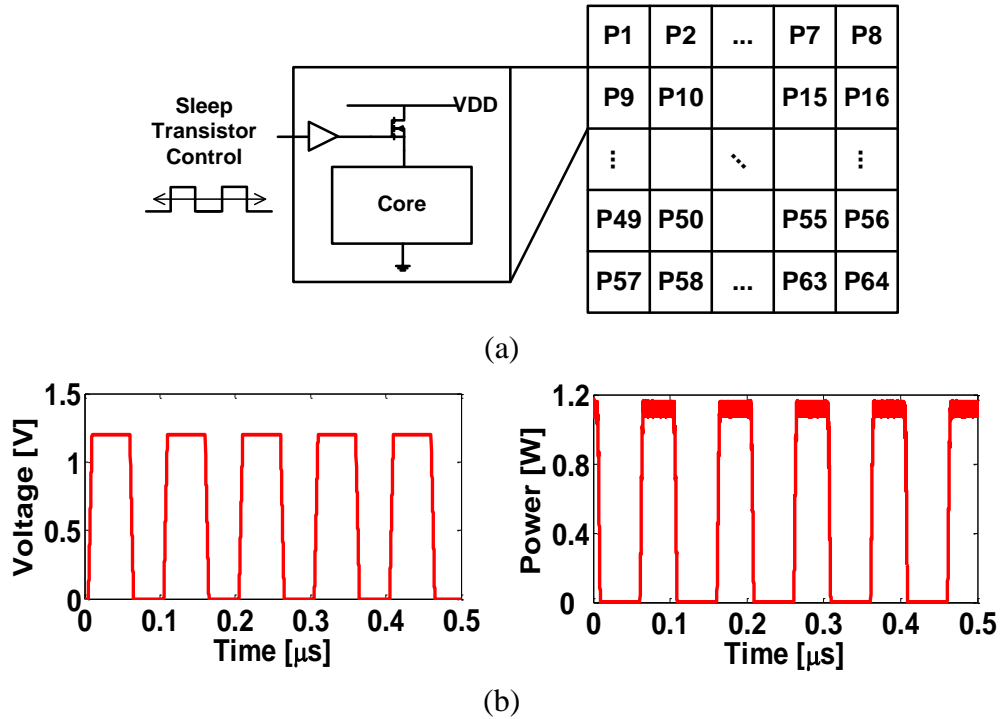


Fig. 64: Core gating based approach to power spectra generation. (a) Sleep transistor signal (5 MHz) and (b) power pattern (5 MHz).

Application Driven Power Spectra: The second approach is to run multiple test applications in individual cores and measure power and temperature to compute the filter response. As power profile generated by each application may not contain significant spectral power at all frequency, we consider average of the filter responses computed using different applications as the extracted filter.

Fig. 65 shows the thermal filter extracted using the practical core-level control closely follows the one from the theoretically ideal small-signal analysis. We observe that the thermal systems behave as the 1st order low-pass filter. The cutoff frequency is located in the low frequency range. Hence, fast time-varying power input has less impact on the temperature while low frequency power variations are more critical. We next study the behavior of the extracted core-to-core cross thermal filters.

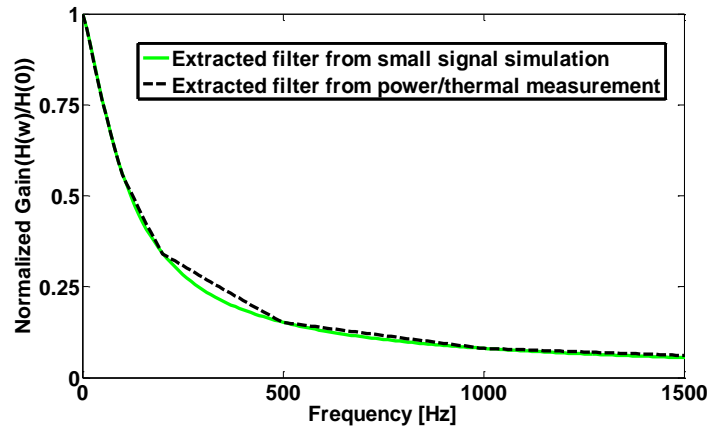


Fig. 65: The thermal filter extraction through small signal simulation (ideal approach for filter extraction) and sleep control based power/thermal measurement. We observe that the proposed practically feasible sleep control approach provides very good accuracy.

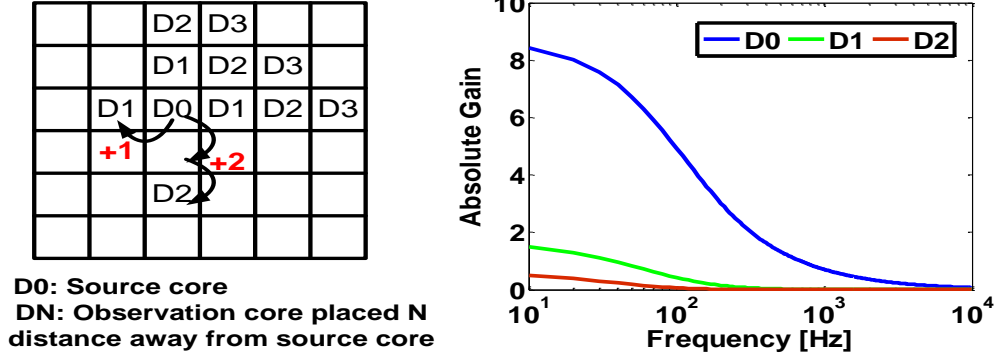


Fig. 66: Filter behavior of thermal system: distance between source core and observation node. We observe that both self (D1) and cross (D1, D2) transfer functions behave as low-pass filter. The strength of the cross transfer function reduces significantly with distance, i.e., power spread in the distant cores will have minimal impact on the temperature of a core. We also see that the effect of cross-transfer function is even less pronounced at a higher frequency.

Fig. 66 shows frequency responses for different location of interest when a power source is applied at core D0. We observe that the gain at the observation point continues to decrease in all frequency range as it moves away from the source. The decrease in the gain due to spatial effect is larger at higher frequencies, i.e., fast varying power source has less impact on neighboring regions. We further note that the filter response between a source and an observation node depends on the physical property of the material system that determines the heat flow. It is independent of the magnitude of the generated power, floorplan of the chip, and architecture. The latter factors modulate the power profile and hence, temperature profile, but not the filter response.

7.5.3 Accuracy of TSI-based Thermal Prediction

We verify the accuracy of the post-silicon TSI-based thermal models against the distributed-RC-based thermal simulator described in Section 7.4. We first create several (60) workloads by randomly assigning the power traces of different application (0.5s of real time data) to different cores and use them for thermal analysis. The same patterns were also run through the baseline distributed-RC-based thermal simulator. Fig. 67 compares the transient temperature variation for

a typical core, and Fig. 68 compares examples of spatial thermal maps at different time instants generated from distributed-RC-based simulation and the proposed approach (with power/thermal measurement driven filter). It can be observed that both transient and spatial temperature variation are well captured.

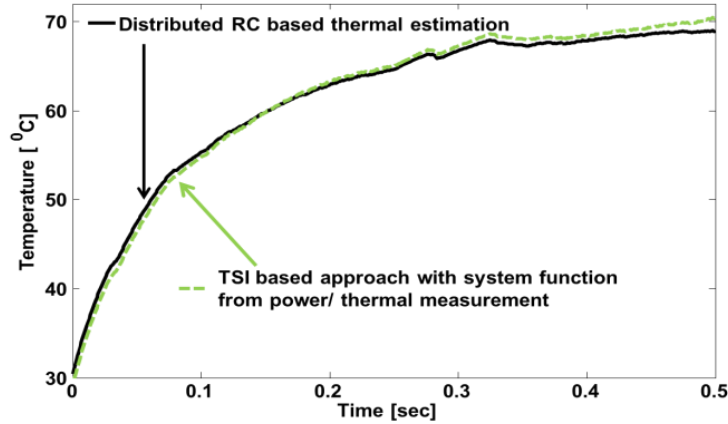


Fig. 67: Estimation error in transient variation of temperature for a typical core in the 64-core system. The simulations were performed considering random workloads created for all 64 cores using random assignments of benchmark applications for SPEC2006 suites.

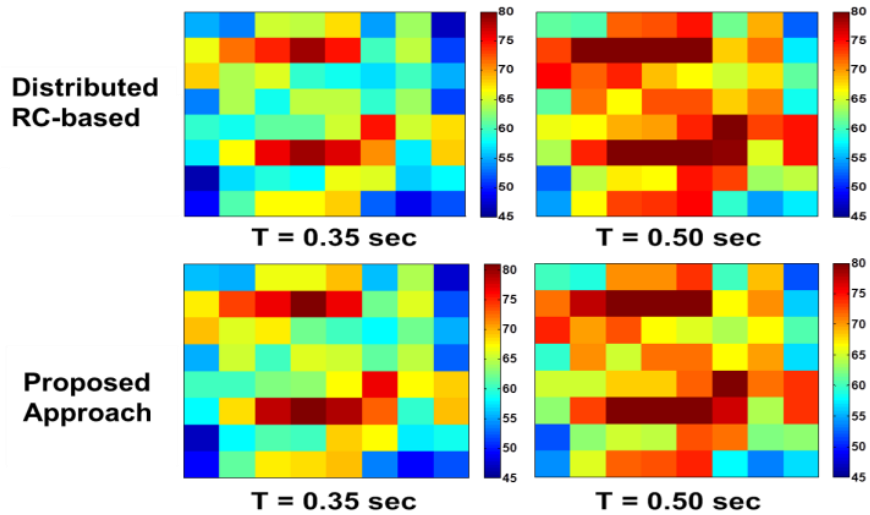


Fig. 68: Estimation error in instances of the spatial thermal field at different time points. The entire simulations consist of random workloads for running in all 64-cores for 500ms. The proposed method represents the estimation using system function extracted from power/thermal measurement. We observe that the proposed TSI-based approach can successfully predict the spatial thermal field.

Fig. 69(a) shows the distribution of estimation errors considering temperature of all cores, all

time points, and all random patterns. We observe that average error is less than 1°C between detailed RC-based thermal simulation (SPICE) and the proposed TSI-based model (the ‘system function from power/thermal measurement’). We study the effect of sensor errors (assuming uniform random error in the range of -3°C to $+3^{\circ}\text{C}$) during TSI on temperature prediction. Fig. 69(a) shows that, even with relatively large measurement error during TSI, the average prediction error increases marginally. Fig. 69(b) shows that the proposed method can accurately predicts the maximum chip temperature as well for different workload.

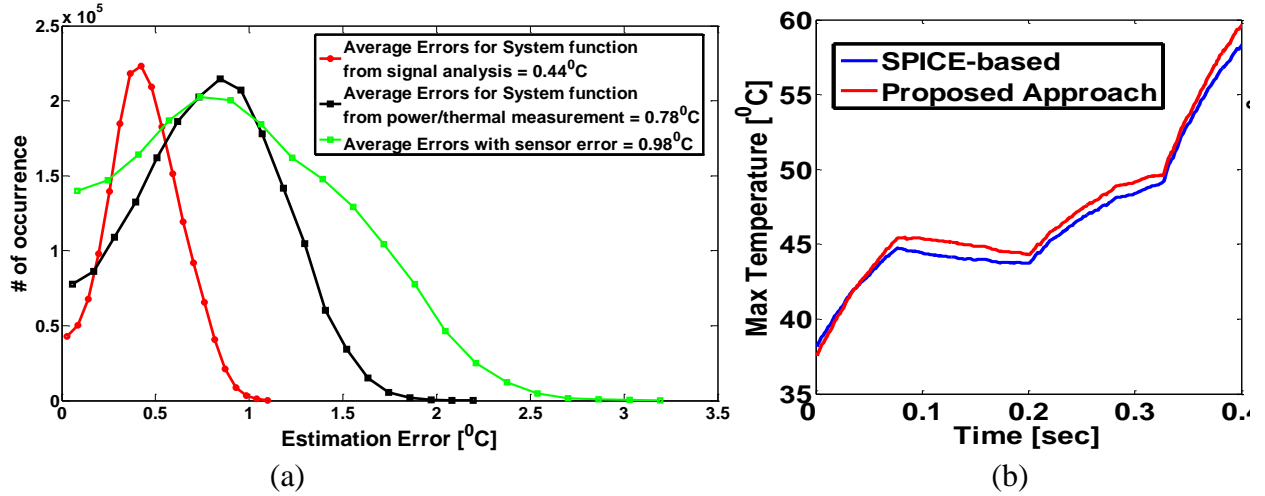


Fig. 69: Accuracy of the proposed approach considering random workload: (a) core-level error statistics considering 64 cores and 60 random workload and (b) prediction error in the peak temperature over time.

7.5.4 Application to Post-Silicon Thermal Prediction

After verifying the accuracy of TSI-based thermal prediction, we next study its effectiveness in post-silicon thermal prediction. We study the ability of TSI in predicting the effect of variations in process corners and thermal conductivity. In this analysis, low-Vt implies a negative 100mV V_{th} shifts for all devices in a chip while high-Vt implies positive 100mV V_{th} shifts. The low-Vth dies have much higher leakage and stronger leakage temperature interaction. Fig. 70(a) and 71(a) show that proposed method captures the effect of chip-to-chip variations in

leakage and thermal conductivity on the extracted thermal filters. We observe that low-Vt die and lower conductivity material (TIM/spreader) increase the gain in the low-frequency range of the filter transfer function. To illustrate the impact of these variations in filter response, we consider Normal random die-to-die variation of V_{th} . Each V_{th} point generated from this Normal distribution represents a unique die for the same many-core processor. For each of such die we consider three different thermal conductivities. TSI is next used to extract the thermal system for all of these die/package condition. The extracted filters for each such instance of the packaged dies are unique. The same workload pattern is applied to all such unique thermal systems to study the effect of process and thermal conductivity variation on-chip temperature. Fig. 70(b) and 71(b) show time-domain temperature variation for a typical core for a chip running the same workload but moved to different V_{th} and thermal conductivity corners. The effect is much stronger when thermal conductivity is low (Fig. 72). This clearly shows the need for post-silicon thermal prediction and the ability of our TSI-based approach. The pre-silicon estimates would have predicted same maximum temperature for all such instances.

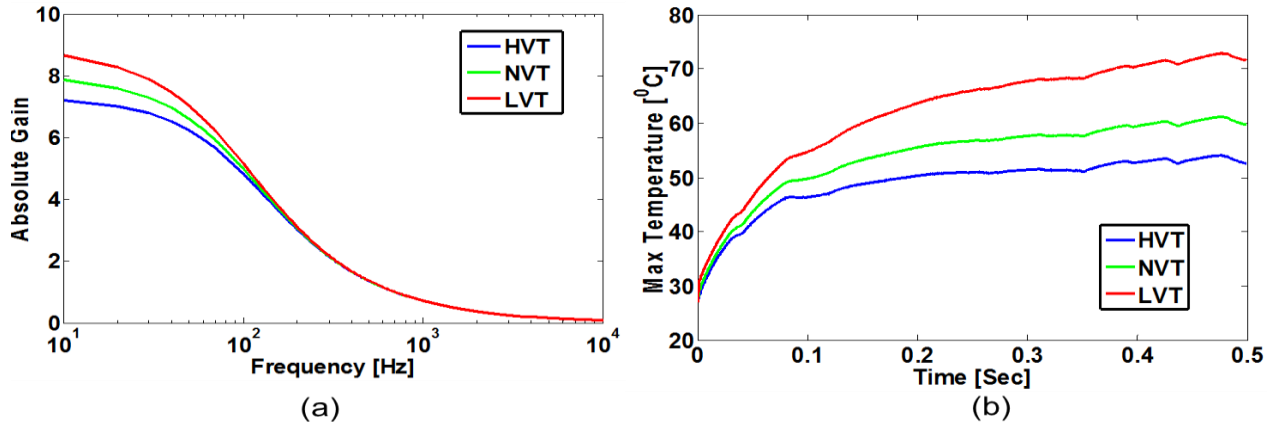


Fig. 70: The application of TSI-based approach on the prediction of impact of process variation on transient temperature. The effect of leakage-temperature interaction is captured in the extracted filter (part-a) showing a higher gain for low-Vt process corners compared to high-Vt corners. Hence, for same workload and dynamic power pattern we observe higher temperature for low-Vt chips than high-Vt chips (part-b).

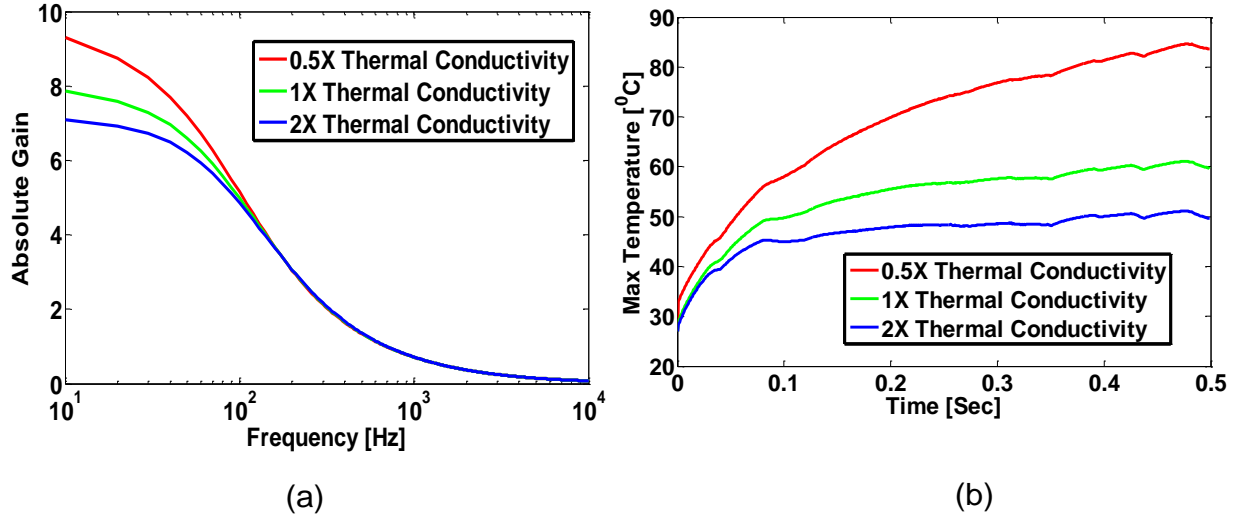


Fig. 71: The application of TSI based approach on the prediction of impact of the conductivity of thermal stack (TIM, spreader, and heat sink) variation on transient temperature. The effect of thermal conductivity is captured in the extracted filter (part-a) showing a higher gain for low-conductivity. Hence, for same workload and dynamic power pattern we observe higher temperature for chips with lower conductivity thermal stack (part-b). The process corner was kept constant in the three simulations.

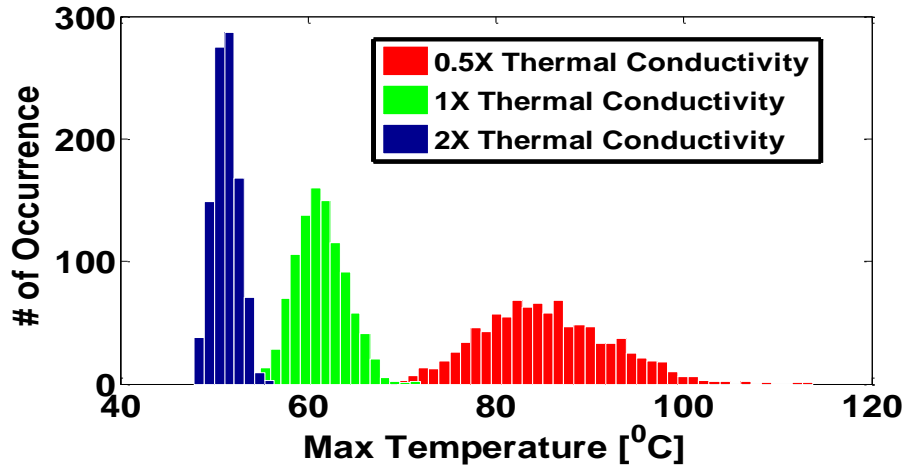


Fig.72: Post-silicon thermal characterization and prediction with chip-to-chip variation in leakage and thermal conductivity: statistical variation in peak temperature of the 64 core processor for same power pattern due to die-to-die leakage variations and different packaging conditions

7.6 Validation through Test-chip

We have designed a test-chip including a poly-resistor-based-heater and thermal sensors to verify the proposed TSI approach through the hardware. The test-chip contains three BJT-based

thermal sensors and a poly-resistor-based heater which mimics a relatively large digital system block. The test-chip has been fabricated in 130nm CMOS technology and mounted on a FR4 board for measurement. Fig. 73 shows the die-photo and the organization of the test-chip. Extraction is performed based on the power/temperature measurement using an oscilloscope, and temperature prediction is performed using MATLAB on a personnel computer with the extracted thermal system.

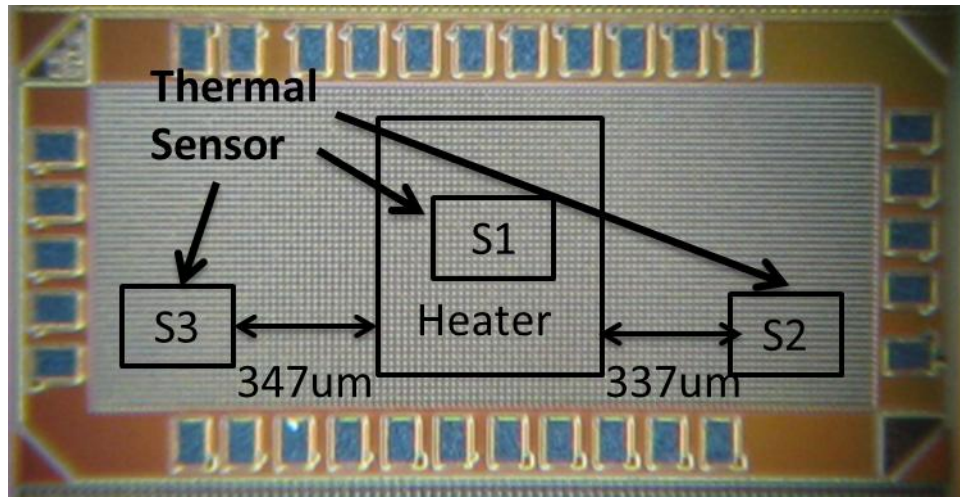
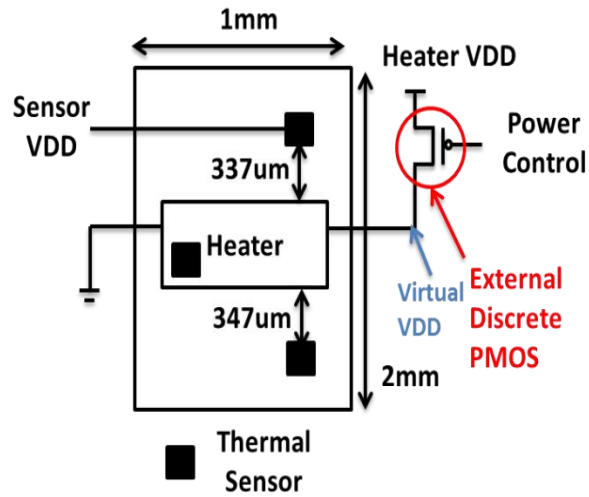


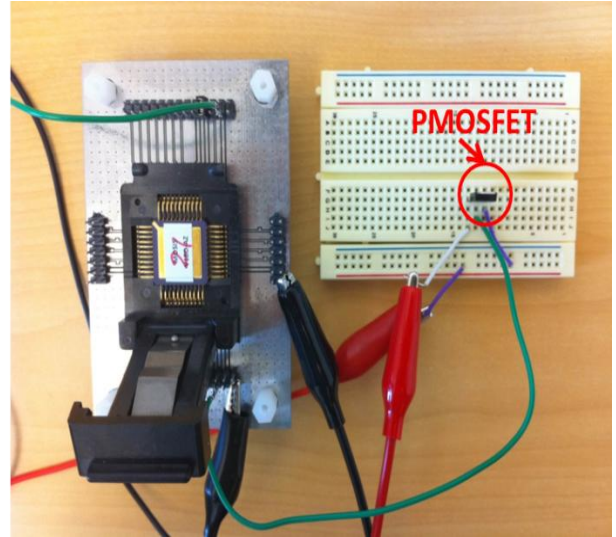
Fig. 73: The die photography.

7.6.1. Implementation

To mimic a maximum operating power/thermal condition of $130\text{W}/\text{cm}^2$, a resistor of 50Ω was used as discussed in Section 6.3. To measure the chip temperature, a BJT-based thermal sensor was implemented as discussed in Section 6.2. As mentioned at Section 6.4, the chip power is controlled externally by PMOS (FQD11P06 from Fairchild semiconductor) to generate a specific frequency signal as shown in Fig. 74. One thermal sensor (S1) is located in the center of the heater, and others (S2, S3) are located $\sim 350\mu\text{m}$ away from the center of the heater as shown in Fig. 73.



(a)



(b)

Fig. 74: Experiment hardware setup: (a) overall schematic and (b) test structure photo.

7.6.2. Measurement Results

Fig. 75 shows the oscilloscope waveforms of thermal sensors and power which is calculated from the measured virtual VDD voltage. Virtual VDD voltage, which is the voltage across the poly-resistor-based heater, is measured at the drain of the sleep transistor through an external pin. The gate of sleep transistor were controlled to turn it on and off at various frequency during the filter extraction period. Although the applied voltage was sinusoid, due to the switching characteristics of a MOSFET, the generated power signal has the shape of a square waveform as shown in Fig. 75.

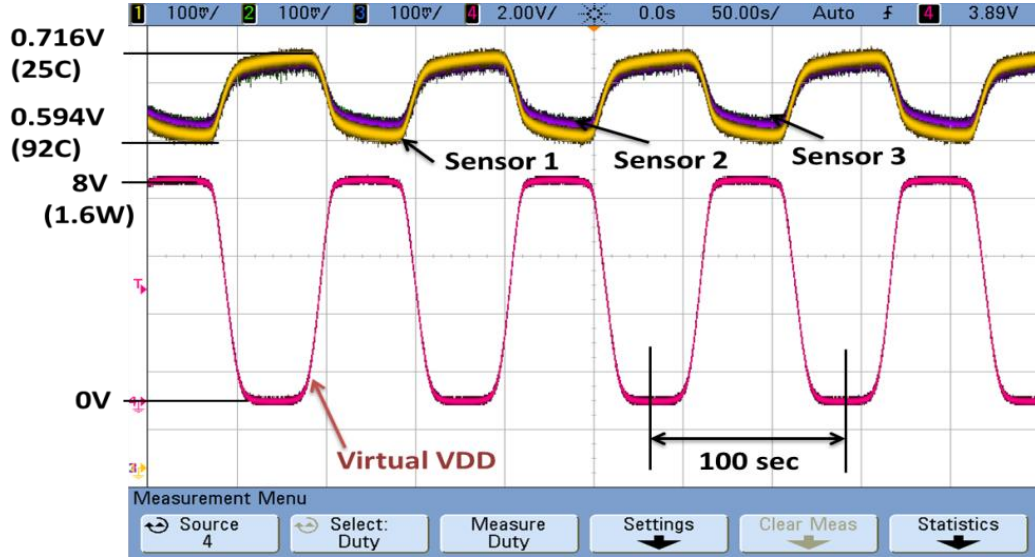


Fig. 75: Oscilloscope waveform of thermal sensor and power.

A. Thermal Filter Extraction

The thermal filter is extracted through the power and temperature measurement data as discussed at Section 7.5.2. To obtain each frequency component of the thermal filter, we apply multiple power patterns with different fundamental frequencies (0.01Hz, 0.1Hz, 1Hz, 10Hz, 100Hz, 1KHz, and 10KHz) to the gate of the sleep transistor. Fig. 76(a),(b) show the measurement of the generated 1Hz power pattern and the associated temperature pattern from the oscilloscope. Fourier transforms were performed on the measured time domain power and temperature patterns. The power spectra converted from 1Hz power signal shows dominant 1Hz frequency component as shown in Fig. 76(c),(d).

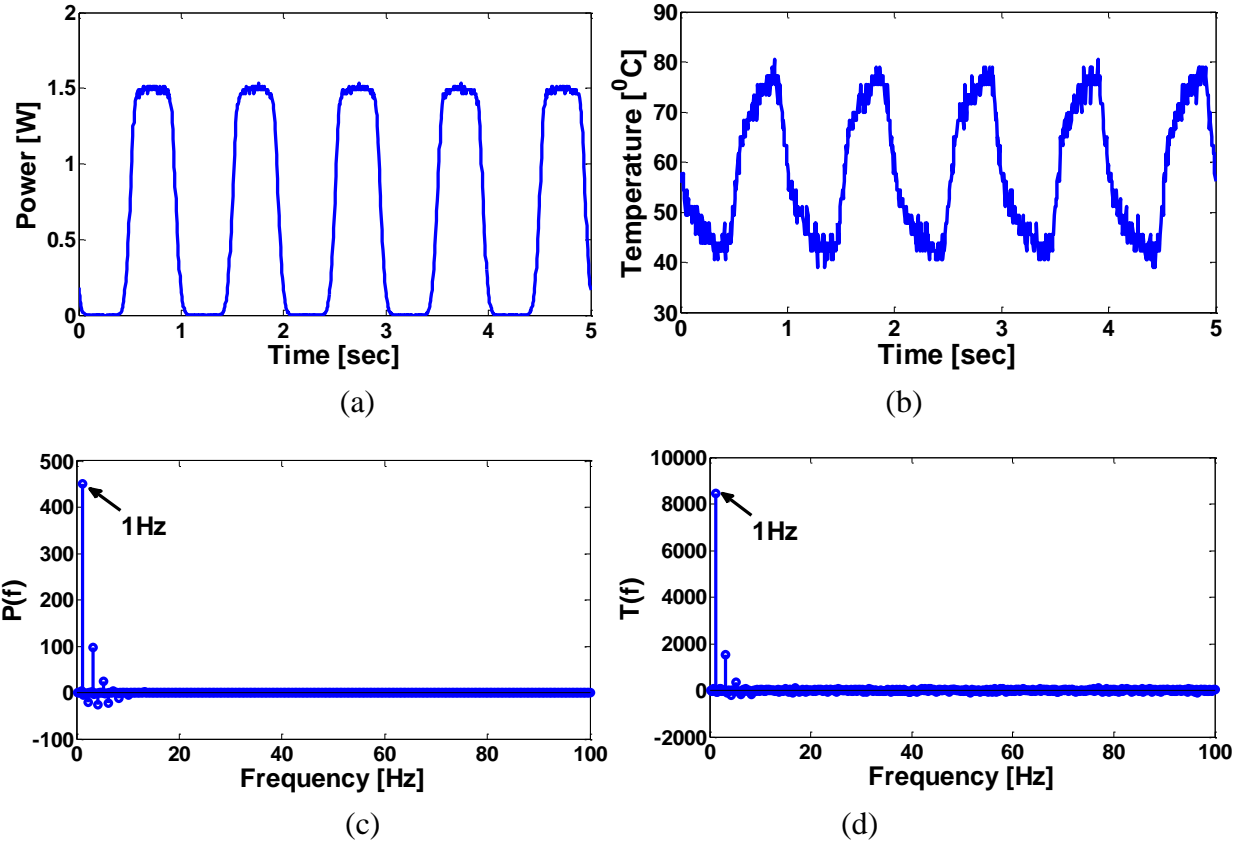


Fig. 76: Thermal system extraction: (a) measured power signal, (b) measured temperature signal, (c) power signal spectra, and (d) temperature signal spectra.

Each fundamental frequency component of the thermal filter is estimated, and intermediate frequency data points are interpolated. This is because extracting all frequency components requires higher extraction time. In addition to that, extracting a fundamental frequency at a time can filter out thermal sensor noise components which tend to be the undesired frequency signal. Fig. 77(a) shows the frequency response of the thermal filters at three different locations. Sensor 1(S1 in Fig. 73), which is located at the center of the heater, has higher absolute gain than the other filters. Other two filters [Sensor 2(S2 in Fig. 73), Sensor 3(S3 in Fig. 73)], which are $\sim 300\mu\text{m}$ away from the heater, have very similar frequency response. This is because two filters are placed similarly in terms of the distance between a thermal sensor and a heater. To evaluate the impact of process/package variation on the thermal filter, multiple (10) chips are measured.

Fig. 77 (b),(c),(d) show the frequency responses of three different thermal filters. We observe the absolute gain in low frequency has more variation than in high frequency. Since the variation in the gain of low frequency for each chip exists, thermal filter extraction for each chip needs to be performed carefully to accurately estimate a temperature.

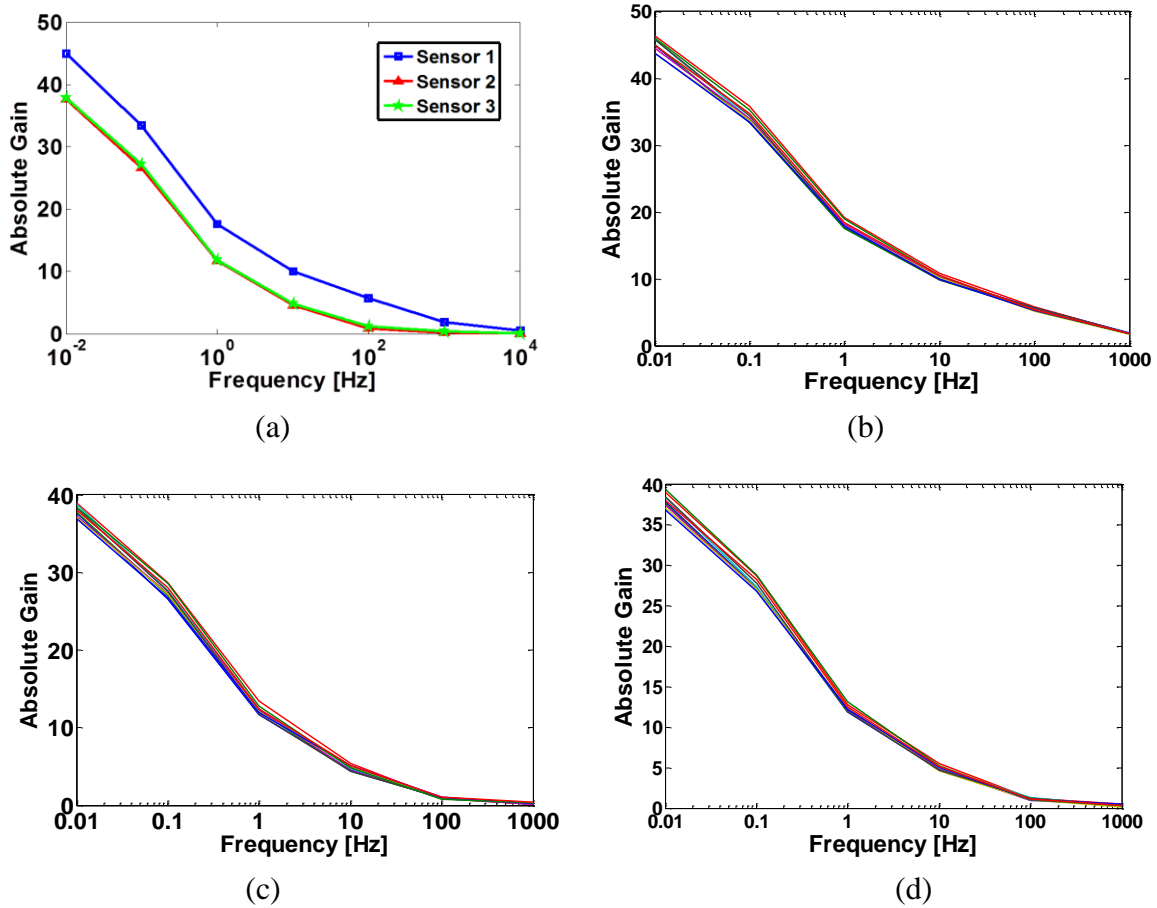
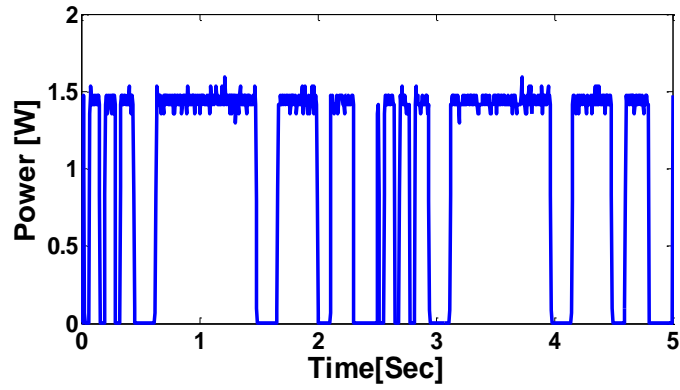


Fig. 77: Extracted thermal filters: (a) three different locations, (b) thermal filter variation for 10 chips (Sensor 1), (c) thermal filter variation for 10 chips (Sensor 2), and (d) thermal filter variation for 10 chips (Sensor 3).

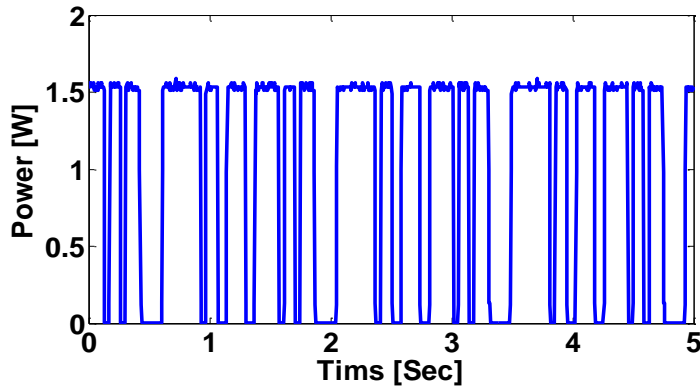
B. Temperature Prediction from Arbitrary Power Variation

We generate exemplary power profiles which include different frequency components as shown in Fig. 78 (a), (b). The power profiles are generated by controlling a sleep transistor signal fed by Keithly power supply. Sleep transistor signal patterns are programmed through ExpressTSP[106] which is an in-built program in the power supply. While the power is applied

to the heater, three temperature sensors measure temperature as shown in Fig. 79. Sensor 1 shows higher temperature than other two sensors for two power patterns because of the spatial effect. Fig. 79 compares the predicted temperature using the proposed approach with the measurement collected by three thermal sensors. Predicted temperatures for different patterns match well to the measurement data from three thermal sensors. Fig. 80 shows predicted temperature waveforms for three different sensors with two different power patterns. As shown in Fig. 80, the average error tends to be overestimated due to the temperature sensor error. Table 5 shows the average estimated errors for Sensor 1, 2, 3 with three different power profiles, which is $\sim 2.3^{\circ}\text{C}$. The proposed approach successfully predicts the temperature which matches the measurement data.

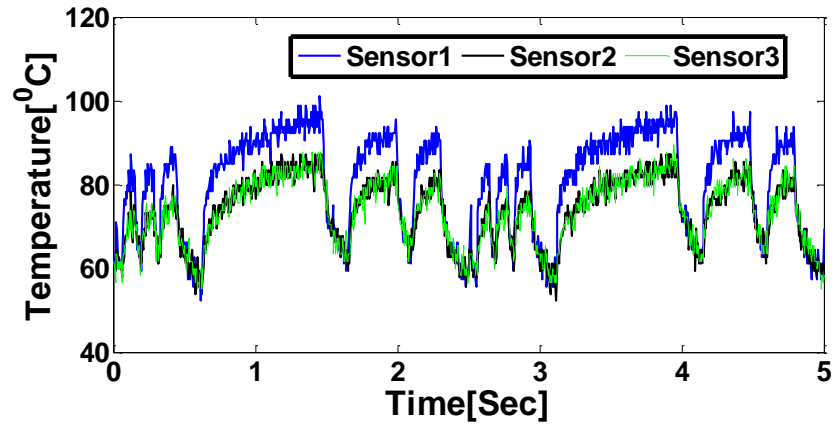


(a)

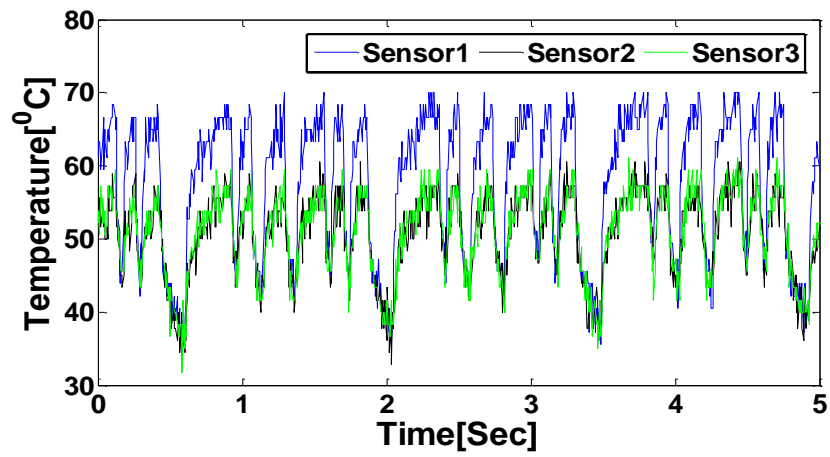


(b)

Fig. 78: Hardware measurement: (a) input power pattern 1 and (b) input power pattern 2.

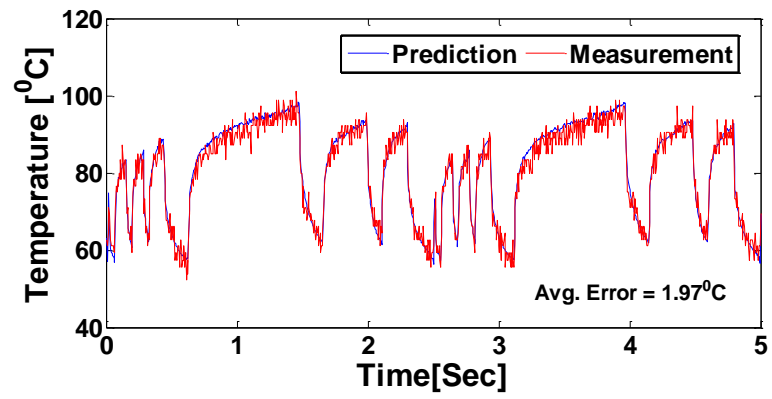


(a)

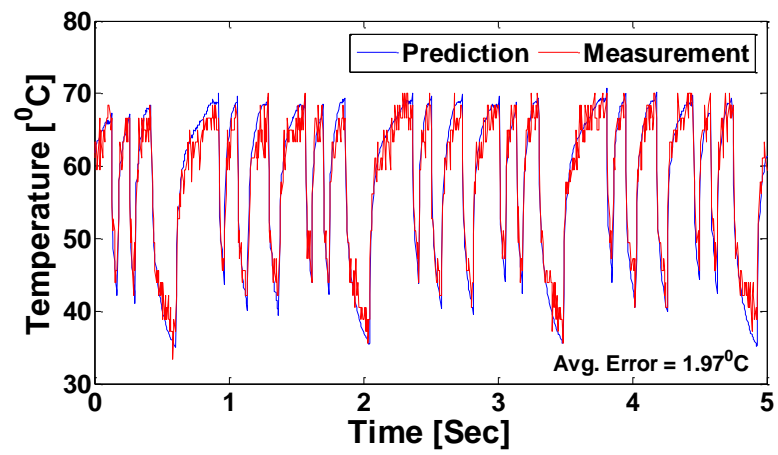


(b)

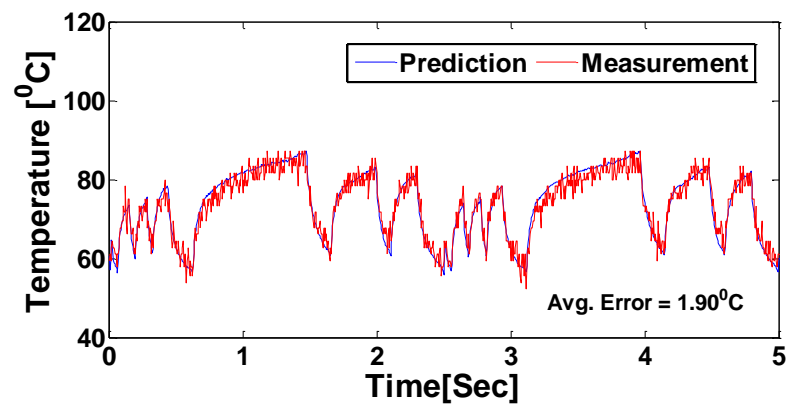
Fig. 79: Hardware measurement: (a) temperature measurement for power pattern 1 and (b) for power pattern 2.



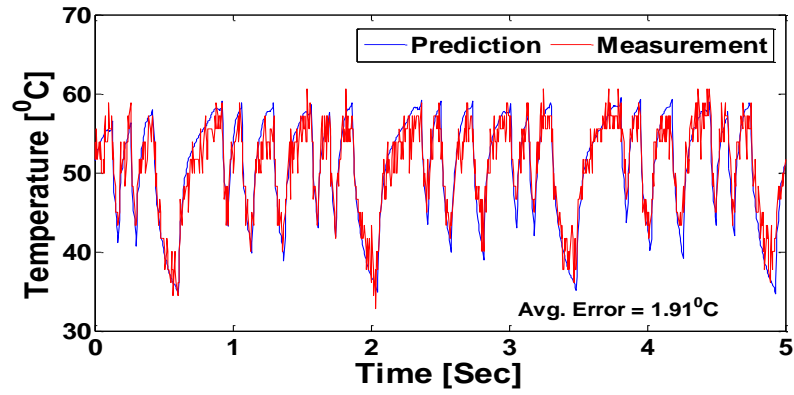
(a)



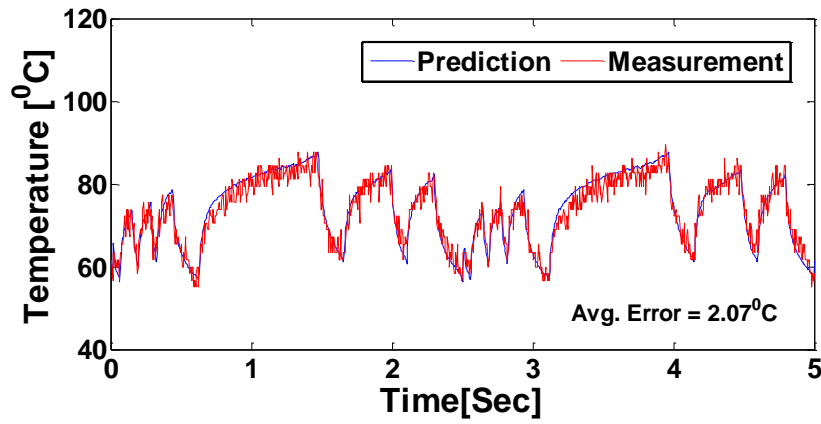
(b)



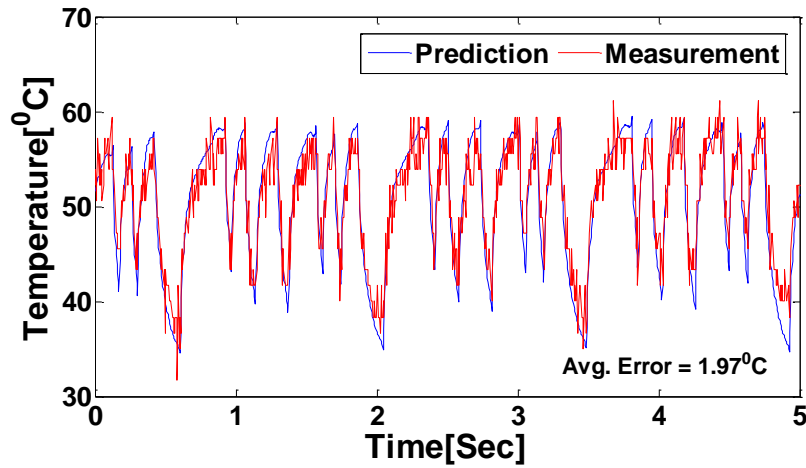
(c)



(d)



(e)



(f)

Fig. 80: Temperature prediction: (a) for Sensor 1 with pattern 1, (b) with pattern 2, (c) for Sensor 2 with pattern 1, (d) with pattern 2, (e) for Sensor 3 with pattern 1, (f) with pattern 2.

Table 5: Average error for 3 power patterns

	Sensor 1	Sensor 2	Sensor 3
Average Error ($^{\circ}\text{C}$)	2.35	2.32	2.31

C. Application of the Digital Sensor

Recent processors employ both an analog and digital sensor. The analog sensors are placed at the temperature-wise critical spot while digital sensors are used for obtaining more temperature information at the less critical spots. Many other applications only use the digital sensors due to the low cost. Therefore, identifying the thermal system through only digital sensors needs to be considered. Along with BJT-based analog thermal sensors, we also implement the digital sensors in the test-chip. The digital sensor is implemented with inverter-based ring oscillators and a counter. Temperature of the digital sensor is estimated based on the frequency of the ring oscillator. Fig. 81(a) shows the good correlation between the digital and analog sensor. The thermal filter extracted from the digital sensor also closely matches to the filter extracted from the analog sensor as shown in Fig. 81(b). Fig. 81(c), (d) show the error estimated from the proposed TSI method from the analog and digital sensor.

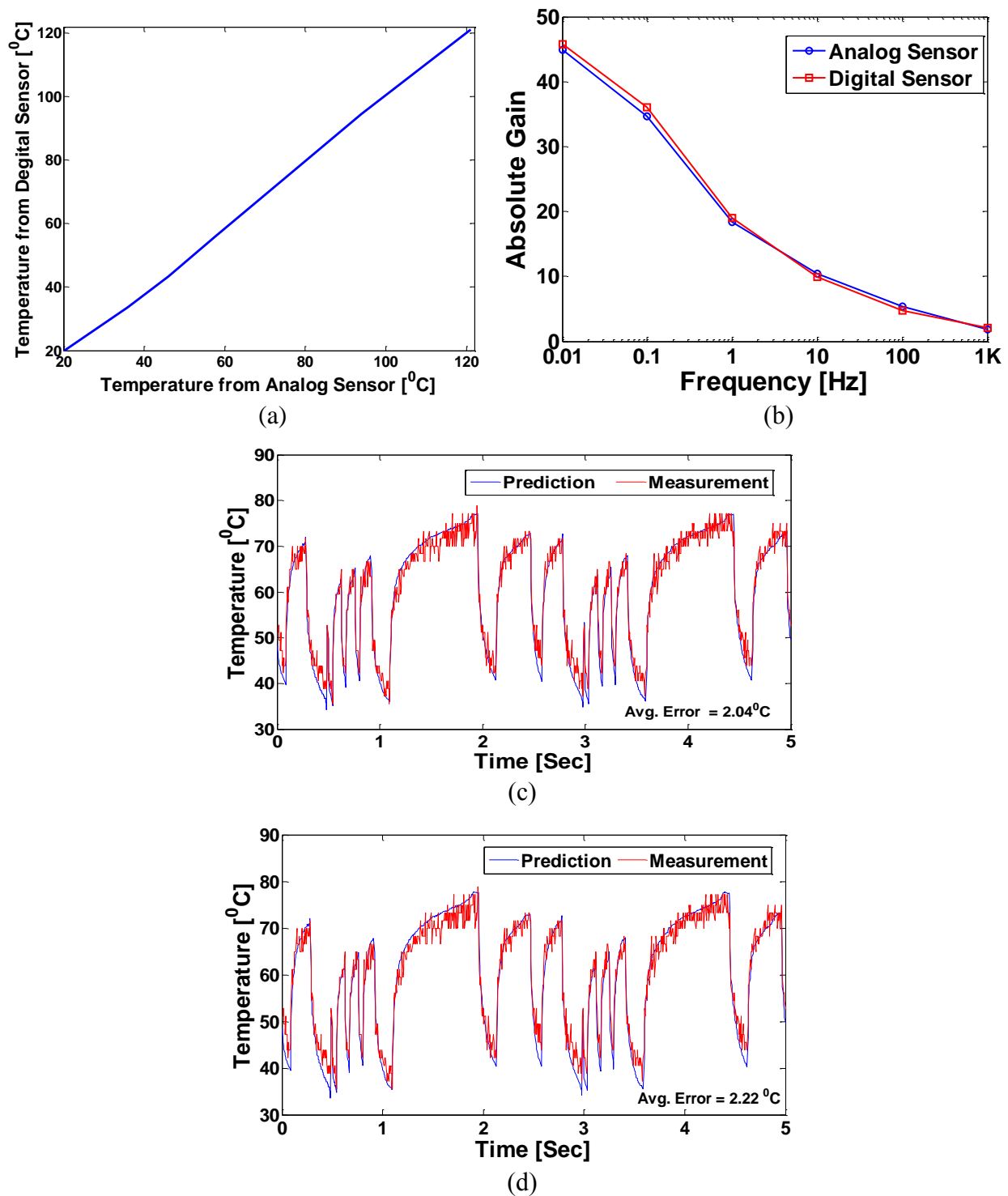


Fig. 81: The Application of a digital sensor: (a) the correlation between the digital and analog sensor, (b) comparison thermal filter extraction from analog and digital sensor, (c) error estimation with an analog sensor, and (d) error estimation with a digital sensor.

D. Phase Behavior of the Thermal system

Fig. 82 shows the phase characteristics of thermal system which is measured from Sensor 1. As expected, as the frequency increases, the frequency response decreases. Assuming this system is the 1st order low-pass filter, the cutoff frequency is $\sim 500\text{Hz}$. According to the phase characteristics, the temperature due to the power input is delayed by the amount of time constant which is the reciprocal of the cutoff frequency. The measurement-based time constant considering process variation can be used to design a temperature sensor and develop more efficient thermal management.

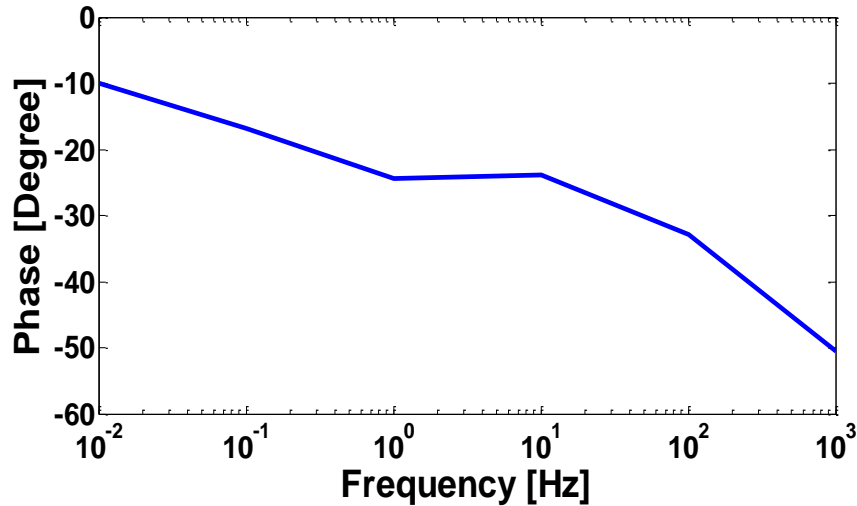


Fig. 82: Phase Characteristic of thermal systems.

7.7. Hardware-based In-Situ Thermal Predictor

The TSI-based thermal modeling provides a unique approach for hardware-based in-situ thermal prediction. One can design an IIR or FIR filter to emulate the low-pass filter nature of thermal system. The filter characteristics can be adapted post-silicon using TSI. This eliminates the need to transform between time and frequency domain during thermal prediction. The time-domain power profile is the input to this in-situ filter and the filter output represents the expected

temperature in time domain. We have verified the feasibility of this approach by implementing an in-situ thermal estimator with 2nd order Chebyshev IIR filters. Fig. 83 shows the estimation error for a representative core is less than $\sim 1.5^{\circ}\text{C}$. The in-situ approach is suitable for on-line prediction to account for time-dependent degradation of thermal property [73].

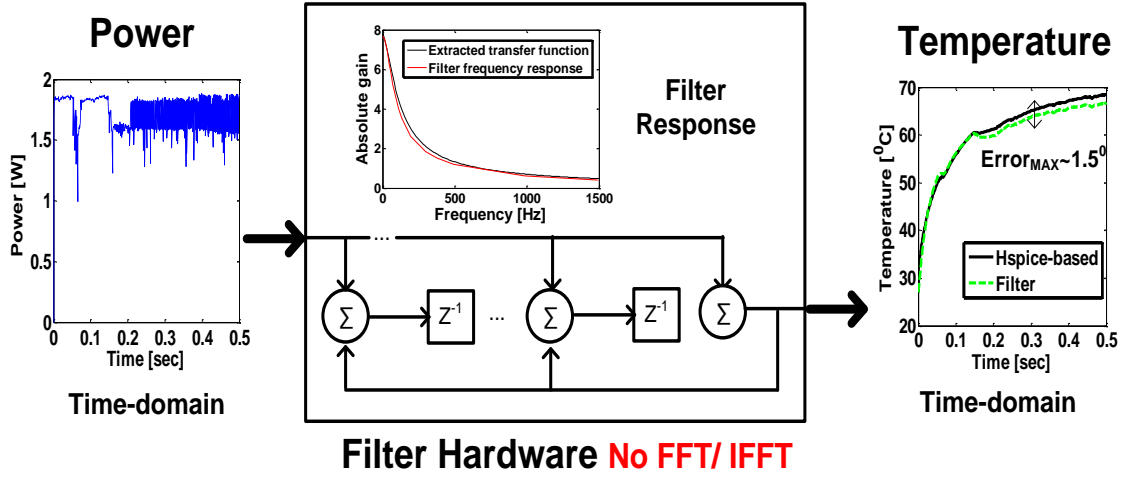


Fig.83: Estimation Error for a representative core for a filter-based implementation (2nd order IIR Chebyshev filter).

7.8. Conclusion

We have presented a methodology for post-silicon thermal prediction. The proposed method first identifies the frequency domain response of the thermal system of a packaged die. The extracted filter is used that for fast chip-specific analysis of transient thermal field considering leakage-temperature feedback. The capabilities of post-silicon characterization of the thermal system can benefit thermal design and management at chip as well as large system level. The test-chip successfully demonstrates the effectiveness of the proposed approach. Estimation error using the proposed approach from the measurement is $\sim 2.3^{\circ}\text{C}$.

CHAPTER VIII

CONCLUSIONS

8.1 Summary and Contribution

For last few decades, device scaling and introduction of new process technologies have played an important role in increasing performance and reliability of digital systems. However, as the feature size scaled down to sub-micrometer domain, electrical engineers started to face many challenges in the process and design. The design methodologies to characterize and compensate for the impact of process/temperature variations on the digital systems need to be considered. This dissertation focuses on developing the methodologies to characterize and compensate for the impact of process and temperature variation across many applications. This chapter presents the summary and contribution and future research we suggest for improving reliability and performance in digital systems.

The summary and contributions of the dissertation present as follows.

In Chapter 3, we presented a design methodology to compensate for process variation in low-power multimedia applications. Parametric failures due to manufacturing variation have been studied with a conventional 6T SRAM. A dynamically reconfigurable SRAM architecture using spatial voltage scaling for low-power mobile multimedia applications is discussed.

In addition to process variation due to the device scaling as shown in Chapter 3, process variation in new technology such as 3D IC technology becomes critical to improve the performance of digital systems. In Chapter 4, we presented a design methodology to characterize

and compensate for process variation in 3D ICs. Pre-bond and post-bond test structure which can test TSV defect is presented, and if necessary, the proposed structure can reconfigure it as a recovery circuit. The effectiveness through standalone simulation as well as a full-chip physical design of a 3D IC has been demonstrated.

To improve the performance and reliability of digital systems, temperature variation as well as process variation needs to be characterized and compensated. In Chapter 5, we presented the design methodology to characterize and compensate for temperature variation in many-core systems. Adaptive spatiotemporal power migration (ASTPM) for burn-in test and run-time operation of many-core chips is presented. ASTPM prevents thermal runaway, improves the quality of test, reduces burn-in test time during the test, and improves the performance during run-time.

In Chapter 6, we presented the challenge of a temperature sensor design. Most methods to deal with thermal management assume accurate thermal sensors are in the chip and normal temperature dependence (i.e., as temperature increases, the frequency decreases.) is affected across the chip. However, power-gating and DVS using multiple supply voltage levels in a chip can cause multiple temperature dependences in one chip, which needs the characterization of inverse temperature dependence in designing the digital thermal sensor. Since the need for a digital sensor increases, the characterization of the digital sensor provides intuition for designing a sensor with better resolution and linearity. With a built-in-poly-resistor-based heater, the characterization of the ITD in digital circuits is presented. Measurement data fabricated in a 130nm CMOS technology is presented for the ITD and sensor characterization.

In Chapter 7, we presented post-silicon system identification and transient thermal prediction. The existing methods to compensate for variations in temperature including the ASTPM are designed and the effectiveness of those methods are verified based on the pre-characterized thermal systems. However, the characteristics of thermal systems vary due to the manufacturing imperfection, which can cause variations in the effectiveness of methods. The presented thermal system identification method captures the characteristics of the thermal systems considering variations in thermal properties. Post-silicon thermal prediction using the extracted thermal system transfer function to predict the transient thermal field a many-core package is presented. Positive feedback path for leakage-temperature interaction considering process variations is implemented. The part of the methodology is verified through the test-chip measurement.

In conclusion, the dissertation presents design methodologies to characterize and compensate for temperature and process variation in low power application to many-core processor systems. The methods to characterize the process or temperature variation allow improving the yield and helping better compensation for the impacts of those variations on the systems. The presented compensation methods also improve the reliability and yield of the systems. The circuit structure presented to test or compensate for variations can apply to the future process to improve the reliability.

8.2 Future Research

Future research related to this dissertation can be as follows. In Section 3, we present a reconfigurable SRAM structure in multimedia applications to compensate for process variation and the effectiveness is verified with exemplary BMP images. A recent mobile application displays many kinds of image types and processes them in a device with small budget of power.

Other image formats have been developed to pursue better quality image with small size data to save a power. Thus, extending the proposed reconfigurable SRAM architecture to different formats of images and video codecs such as JPEG, JPEG2000, and H.264 codec would be meaningful and practical for compensating the impact of process variation.

In Section 4, we present the characterization and compensation for process variation especially in 3D ICs. We have primarily focused on TSV-short defect in this dissertation. Based on the literature survey, TSV-short is very critical in the functionality of 3D ICs. However, as the technology is matured and developed, other defects such as TSV-to-TSV coupling defect can become more critical in the functionality of 3D ICs. To fully characterize and compensate for new technology process variation, new or modified test structure to consider other defects will be necessary.

Regarding the thermal system identification study which is mentioned in Section 6, the thermal filter extraction and verification with different packages and cooling conditions such as TEC or microfluidic cooling would be very interesting and helpful for better understanding of thermal management. Especially, in a 3D IC, complicated physical structure requires more effort for modeling the thermal component and simulation. On top of the difficulty of modeling and simulation, the variation in thermal properties due to the imperfect manufacturing can cause unexpected temperature profiles, which can fail thermal management. Therefore, characterizing the thermal system with different cooling solutions with the TSI will help to develop more efficient thermal management policy. Hardware validation on the TSI demonstrated in this thesis the case with one power source and multiple observations, which is the part of the TSI method. Therefore, the verification for the case with multiple power sources and multiple observations with a test-chip needs to be performed to complete the hardware validation on the TSI method.

APPENDIX A

DERIVATIONS OF THE ZTC POINT IN LINEAR AND SATURATION REGIONS

In this appendix, detailed computations of the ZTC point in the linear and saturation region, which is used in Chapter VII, are presented. The ZTC V_G point is a voltage level where a drain current remain same over temperature, which can be calculated by solving for $\frac{\partial I_D}{\partial T} = 0$.

- Derivation of the ZTC point in the linear region

I_D in the linear is given as follows:

$$I_D = \frac{W}{L} C_{OX} \mu_n(T) \left\{ (V_{GS} - V_{th}(T)) V_{DS} - \frac{m}{2} V_{DS}^2 \right\} \quad (30)$$

where μ_n is the effective channel mobility, C_{OX} is the gate oxide capacitance per unit area, W and L are the channel width and length; respectively, and m is the body effect coefficient. The temperature dependence of mobility and V_{th} [102, 103] is respectively expressed by (31) and (32):

$$\mu(T) = \mu(T_0) \left(\frac{T}{T_0} \right)^{-\theta} \quad (31)$$

$$V_{th}(T) = V_{th}(T_0) - k(T - T_0) \quad (32)$$

where T_0 is the initial room absolute temperature (300K), θ is the mobility temperature exponent, k is the temperature coefficient of the threshold voltage.

By inserting $\mu(T)$ in (31) and V_{th} in (32) into (30), we get

$$\begin{aligned}
I_D &= \frac{W}{L} C_{OX} \mu_n(T) \left\{ (V_{GS} - V_{th}(T)) V_{DS} - \frac{m}{2} V_{DS}^2 \right\} \\
&= \frac{W}{L} C_{OX} \mu_n(T_0) \left(\frac{T}{T_0} \right)^{-\theta} \left\{ (V_{GS} - V_{th}(T)) V_{DS} - \frac{m}{2} V_{DS}^2 \right\} \\
&= \frac{W}{L} C_{OX} \mu_n(T_0) \left(\frac{T}{T_0} \right)^{-\theta} \left\{ (V_{GS} - (V_{th}(T_0)) - k(T - T_0)) V_{DS} - \frac{m}{2} V_{DS}^2 \right\} \\
&= \underbrace{\frac{W}{L} C_{OX} \mu_n(T_0) T_0^\theta}_S T^{-\theta} \left\{ (V_{GS} - V_{th}(T_0) + k(T - T_0)) V_{DS} - \frac{m}{2} V_{DS}^2 \right\}
\end{aligned} \tag{33}$$

For $\frac{\partial I_D}{\partial T} = 0$, we calculate V_{GS} as follows:

$$\begin{aligned}
\frac{\partial I_D}{\partial T} &= \left[S T^{-\theta} \left\{ (V_{GS} - V_{th}(T_0) + k(T - T_0)) V_{DS} - \frac{m}{2} V_{DS}^2 \right\} \right]' \\
&= S \left[(-\theta) T^{-\theta-1} \left\{ (V_{GS} - V_{th}(T_0) + k(T - T_0)) V_{DS} - \frac{m}{2} V_{DS}^2 \right\} + T^{-\theta} k V_{DS} \right] \\
&= 0 \\
\theta T^{-\theta-1} \left\{ (V_{GS} - V_{th}(T_0) + k(T - T_0)) V_{DS} - \frac{m}{2} V_{DS}^2 \right\} &= T^{-\theta} k V_{DS} \\
\theta T^{-1} \left\{ (V_{GS} - V_{th}(T_0) + k(T - T_0)) V_{DS} - \frac{m}{2} V_{DS}^2 \right\} &= k V_{DS} \\
(V_{GS} - V_{th}(T_0) + k(T - T_0)) V_{DS} - \frac{m}{2} V_{DS}^2 &= \frac{k V_{DS}}{\theta} T \\
V_{GS} - V_{th}(T_0) + k(T - T_0) - \frac{m}{2} V_{DS} &= \frac{k}{\theta} T \\
V_{GS} &= V_{th}(T_0) - k(T - T_0) + \frac{m}{2} V_{DS} + \frac{k}{\theta} T
\end{aligned}$$

- Derivation of the ZTC point in the saturation region

I_D in the saturation regions is given as follows:

$$I_D = \frac{W}{L} C_{OX} \mu_n(T) \frac{(V_{GS} - V_{th}(T))^\alpha}{m} \tag{34}$$

By inserting $u(T)$ in (31) and V_{th} in (32) into (34), we get

$$\begin{aligned}
I_D &= \frac{W}{L} C_{ox} \mu_n(T) \frac{(V_{GS} - V_{th}(T))^\alpha}{m} \\
&= \frac{W}{L} C_{ox} \mu_n(T_0) \left(\frac{T}{T_0}\right)^{-\theta} \frac{(V_{GS} - V_{th}(T_0) + k(T - T_0))^\alpha}{m}
\end{aligned}$$

For $\frac{\partial I_D}{\partial T} = 0$, we calculate V_{GS} as follows:

$$\begin{aligned}
\frac{\partial I_D}{\partial T} &= \underbrace{\frac{W}{L} C_{ox} \mu_n(T_0) T_0^\theta}_{s} T^{-\theta} \frac{(V_{GS} - V_{th}(T_0) + k(T - T_0))^\alpha}{m} \frac{\partial I_D}{\partial T} \\
&= \left[s T^{-\theta} \frac{(V_{GS} - V_{th}(T_0) + k(T - T_0))^\alpha}{m} \right]' \\
&= \frac{s}{m} [(-\theta) T^{-\theta-1} (V_{GS} - V_{th}(T_0) + k(T - T_0))^\alpha + T^{-\theta} \alpha (V_{GS} - V_{th}(T_0) + \\
&\quad k(T - T_0))^{\alpha-1}] = 0
\end{aligned}$$

$$\theta T^{-\theta-1} (V_{GS} - V_{th}(T_0) + k(T - T_0))^\alpha = T^{-\theta} \alpha (V_{GS} - V_{th}(T_0) + k(T - T_0))^{\alpha-1}$$

$$\theta T^{-1} = \alpha (V_{GS} - V_{th}(T_0) + k(T - T_0))^{-1}$$

$$V_{GS} = V_{th}(T_0) - k(T - T_0) + \frac{\alpha T}{\theta}$$

REFERENCES

- [1] T. Austin, D. Blaauw, S. Mahlke, T. Mudge, C. Chakrabarti, and W. Wolf, "Mobile supercomputers," *Computer*, vol. 37, no. 5, pp. 81- 83, May 2004.
- [2] S. Yang, W. Wolf, and N. Vijaykrishnan, "Power and performance analysis of motion estimation based on hardware and software realizations," *IEEE Transactions on Computers*, vol. 54, no. 6, pp. 714- 726, June 2005.
- [3] F. Catthoor, E. Greef, and S. Suytack, *Custom Memory Management Methodology*, Boston: Kluwer Academic, 1998.
- [4] G. Chen and M. Kandemir, "Optimizing address code generation for array-intensive DSP applications," in *International Symposium on Code generation and optimization*, pp. 141- 152, Mar. 2005.
- [5] K. Choi, K. Dantu, W. Cheng, and M. Pedram, "Frame-based dynamic voltage and frequency scaling for a MPEG decoder," in *IEEE/ACM International Conference on Computer Aided Design*, pp. 732- 737, Nov. 2002.
- [6] S. Mukhopadhyay, H. Mahmoodi, and K. Roy, "Modeling of failure probability and statistical design of SRAM array for yield enhancement in nanoscaled CMOS," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 12, pp. 1859- 1880, Dec. 2005.
- [7] K. Agarwal and S. Nassif, "Statistical analysis of SRAM cell stability," in *ACM/IEEE Design Automation Conference*, pp.57-62, 2006.

- [8] R. Kanj, R. Joshi, and S. Nassif, "Mixture importance sampling and its application to the analysis of SRAM designs in the presence of rare failure events," in *ACM/IEEE Design Automation Conference*, pp.69-72, 2006.
- [9] F. Kurdahi, A. Eltawil, A. K. Djahromi, M. Makhzan, and S. Cheng, "Error-Aware Design," in *10th Euromicro Conference on Digital System Design Architectures, Methods and Tools*, pp.8-15, Aug. 2007.
- [10] A. K. Djahromi, A. M. Eltawil, F. Kurdahi, and R. Kanj, "Cross Layer Error Exploitation for Aggressive Voltage Scaling," in *8th International Symposium on Quality Electronic Design*, pp.192-197, Mar. 2007.
- [11] J. George, B. Marr, B. E. S. Akgul, and K. V. Palem, "Probabilistic Arithmetic and Energy Efficient Embedded Signal Processing," in *International Conference on Compilers, Architecture and Synthesis for Embedded Systems*, Oct. 2006.
- [12] S. Cheemavalagu, P. Korkmaz, and K. V. Palem, "Ultra low-energy computing via probabilistic algorithms and devices: CMOS device primitives and the energy-probability relationship," in *International Conference on Solid State Devices and Materials*, pp. 402–403, Sept. 2004.
- [13] K. Yi, S. Y. Cheng, F. Kurdahi, and A. Eltawil, "A partial memory protection scheme for higher effective yield of embedded memory for video data," in *Asia-Pacific Computer Systems Architecture Conference*, pp.1-6, Aug. 2008.
- [14] B. Wicht, T. Nirschl, and D. Schmitt-Landsiedel, "Yield and speed optimization of a latch-type voltage sense amplifier," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 7, pp. 1148- 1158, July 2004.

- [15] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, Apr. 2004.
- [16] K. Zhang, U. Bhattacharya, Z. Chen, F. Hamzaoglu, D. Murray, N. Vallepalli, Y. Wang, B. Zheng, and M. Bohr, "A 3-GHz 70MB SRAM in 65nm CMOS technology with integrated column-based dynamic power supply," in *IEEE International Solid-State Circuits Conference*, pp. 474-611 Vol. 1, Feb. 2005.
- [17] "<http://www.imageprocessingplace.com/>".
- [18] M. Motoyoshi, "Through-Silicon Via (TSV)," *Proceedings of the IEEE*, vol. 97, no. 1, pp. 43-48, Jan. 2009.
- [19] R.S. Patti, "Three-Dimensional Integrated Circuits and the Future of System-on-Chip Designs," *Proceedings of the IEEE*, vol. 94, no. 6, pp. 1214-1224, June 2006.
- [20] E. J. Marinissen and Y. Zorian, "Testing 3D chips containing through-silicon vias", in *International Test Conference*, 2009, pp1-11.
- [21] X. Zhao, D.L. Lewis, H.-H.S Lee, and S. K. Lim, "Pre-bond Testable Low-Power Clock Tree Design for 3D Stacked ICs", in *IEEE International Conference on Computer Aided Design*, 2009, pp. 184-190.
- [22] H. Lee and K. Chakrabarty, "Test Challenges for 3D Integrated Circuits", *IEEE Design and Test of Computers*, vol. 26, no. 5, pp. 26-35, Sept. 2009.
- [23] X. Wu, P. Falkenstern, K. Chakrabarty, and Y. Xie, "Scan-chain design and optimization for three-dimensional integrated circuits", *ACM Journal on Emerging Technologies in Computing Systems*, vol. 5, no. 2, pp. 1-26, July 2009.

- [24] C.Y. Lo, Y. T. Hsing, L. M. Denq, and C. W. Wu, "SOC Test Architecture and Method for 3-D ICs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 29, no. 10, pp. 1645-1649, Oct. 2010.
- [25] E. J. Marinissen, J. Verbree, and M. Konijnenburg, "A structured and scalable test access architecture for TSV-based 3D stacked ICs," in *VLSI Test Symposium*, 2010, pp. 269-274.
- [26] M. Tsai, A. Klooz, A. Leonard, J. Appel, and P. Franzon, "Through Silicon Via(TSV) defect/pinhole self-test circuit for 3D-IC," in *IEEE International Conference 3D System Integration*, 2009, pp. 1-8.
- [27] P. Y. Chen, C. W. Wu, and D. M. Kwai, "On-Chip TSV Testing for 3D IC before Bonding Using Sense Amplification," in *Asian Test Symposium*, 2009, pp. 450-455.
- [28] P. Y. Chen, C. W. Wu, and D. M. Kwai, "On-chip testing of blind and open-sleeve TSVs for 3D IC before bonding," in *VLSI Test Symposium*, 2010, pp. 263-268.
- [29] M. Cho, L. Chang, D. Kim, S. Lim, and S. Mukhopadhyay, "Design Method and Test Structure to Characterize and Repair TSV Defect Induced Signal Degradation in 3D System," in *International Conference on Computer-Aided Design*, 2010, pp. 694-697.
- [30] HSPICE, Synopsys, Inc., Mountain View, CA, 2010.
- [31] Predictive Technology Model (PTM), Available HTTP: <http://www.eas.asu.edu/~ptm/>.
- [32] T. Thorolfsson, K. Gonsalves, and P.D. Franzon, "Design Automation for a 3DIC FFT Processor for Synthetic Aperture Radar: A Case Study," in *Design Automation Conference*, 2009, pp. 51-56.

- [33] D. H. Kim, K. Athikulwongse, and S. K. Lim, "A Study of Through-Silicon-Via Impact on the 3D Stacked IC Layout," in *IEEE International Conference on Computer Aided Design*, 2009, pp. 674-680.
- [34] Y. J. Lee, M. Pathak, C. Liu, M. Jung, and S. K. Lim, "Design and Timing Optimization of a 3D Stacked Microprocessor," in *ACM International Workshop on Timing Issues in the Specification and Synthesis of Digital Systems*, 2010, pp. 19-24.
- [35] R.-P. Vollertsen, "Burn-In," in *IEEE Int'l Integrated Reliability Workshop Final Report*, IEEE Standard Office, pp.167-173, 1999.
- [36] Quality and Reliability, ASIC Products Application Note SA14-2280-03, rev. 3, Microelectronics Division, IBM, 1999.
- [37] JEDEC Solid-State Technology Association, "Early Life Failure Rate Calculation Procedure for Electronic Components," Apr. 2000. Available: JEDEC Std. JESD74.
- [38] S. Sheu and Y. Chien, "Minimizing Cost-Functions Related to Both Burn-In and Field-Operation Under a Generalized Model," *IEEE Transactions on Reliability*, vol. 53, no. 3, pp. 435- 439, Sept. 2004.
- [39] M. F. Zakaria, Z. A. Kassim, M. P. Ooi, and S. Demidenko, "Reducing burn-in time through high-voltage stress test and Weibull statistical analysis," *IEEE Design & Test of Computers*, vol. 23, no. 2, pp. 88- 98, Mar. 2006.
- [40] A. Vassighi, O. Semenov, M. Sachdev, A. Keshavarzi, and C. Hawkins, "CMOS IC technology scaling and its impact on burn-in," *IEEE Transactions on Device and Materials Reliability*, vol. 4, no. 2, pp. 208- 221, June 2004.
- [41] M. Meterelliyoz, H. Mahmoodi, and K. Roy, "A leakage control system for thermal stability during burn-in test," in *International Test Conference*, Nov. 2005.

- [42] P. Nigh and A. Gattiker, "Test method evaluation experiments and data," in *International Test Conference*, pp.454-463, 2000.
- [43] P. Tadayon, "Thermal challenges during microprocessor testing," [Online Document], Available HTTP: <http://download.intel.com/technology/itj/q32000/pdf/thermal.pdf>.
- [44] S. Mukhopadhyay, C. Neau, R. T. Cakici, A. Agarwal, C. H. Kim, and K. Roy, "Gate leakage reduction for scaled devices using transistor stacking," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 11, no. 4, pp. 716-730, Aug. 2003.
- [45] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits," *Proceedings of the IEEE*, vol. 91, no. 2, pp. 305- 327, Feb. 2003.
- [46] J. Tschanz, S. G. Narendra, Y. Ye, B. A. Bloechel, S. Borkar, and V. De, "Dynamic sleep transistor and body bias for active leakage power control of microprocessors," *IEEE Journal of Solid-State Circuits*, vol. 38, no.11, pp. 1838- 1845, Nov. 2003.
- [47] S. Borkar, "Thousand Core Chips—A Technology Perspective," in *ACM/IEEE Design Automation Conference*, pp.746-749, June 2007.
- [48] S. Vangal, J. Howard, G. Ruhl, S. Dighe, H. Wilson, J. Tschanz, D. Finan, P. Iyer, A. Singh, T. Jacob, S. Jain, S. Venkataraman, Y. Hoskote, and N. Borkar, "An 80-Tile 1.28TFLOPS Network-on-Chip in 65nm CMOS," in *IEEE International Solid-State Circuits Conference*, pp. 98-589, Feb. 2007.
- [49] D. C. Sekar, A. Naeemi, R. Sarvari, J. A. Davis, and J. D. Meindl, "Intsim: a CAD tool for optimization of multilevel interconnect networks," in *IEEE/ACM International Conference on Computer-Aided Design*, pp. 560-567, Nov. 2007.

- [50] International Technology Roadmap for Semiconductors (ITRS), Available HTTP: <http://public.itrs.net/>.
- [51] R. J. Ribando and K. Skadron, "Many-core design from a thermal perspective," in *45th ACM/IEEE Design Automation Conference*, pp.746-749, June 2008.
- [52] A. K. Coskun, T. T. Rosing, K. A. Whisnant, and K. C. Gross, "Static and Dynamic Temperature-Aware Scheduling for Multiprocessor SoCs," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 16, no. 9, pp. 1127-1140, Sept. 2008.
- [53] D. Brooks and M. Martonosi, "Dynamic thermal management for high-performance microprocessors," in *International Symposium on High-Performance Computer Architecture*, pp. 171-182, 2001.
- [54] P. Chaparro, J. Gonzalez, G. Magklis, C. Qiong, and A. Gonzalez, "Understanding the Thermal Implications of Multi-Core Architectures," *IEEE Transactions on Parallel and Distributed System*, vol. 18, no. 8, pp. 1055-1065, Aug. 2007.
- [55] R. McGowen, C.A. Poirier, C. Bostak, J. Ignowski, M. Millican, W. H. Parks, and S. Naffziger, "Power and temperature control on a 90-nm Itanium family processor," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 1, pp. 229- 237, Jan. 2006.
- [56] J. Tschanz, N. S. Kim, S. Dighe, J. Howard, G. Ruhl, S. Vanga, S. Narendra, Y. Hoskote, H. Wilson, C. Lam, M. Shuman, C. Tokunaga, D. Somasekhar, S. Tang, D. Finan, T. Karnik, N. Borkar, N. Kurd, and V. De, "Adaptive Frequency and Biasing Techniques for Tolerance to Dynamic Temperature-Voltage Variations and Aging," in *IEEE International Solid-State Circuits Conference*, pp. 292-604, Feb. 2007.

- [57] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers, "The case for lifetime reliability-aware microprocessors," in *International Symposium on Computer Architecture*, pp. 276-287, June 2004.
- [58] K. Sankaranarayanan, S. Velusamy, M.R. Stan, and K. Skadron, "A Case for Thermal-Aware Floorplanning at the Microarchitectural Level," *The Journal of Instruction-Level Parallelism*, vol. 7, Oct. 2005.
- [59] R. Rao, S. Vrudhula, and C. Chakrabarti, "Throughput of multi-core processors under thermal constraints," in *International Symposium on Low Power Electronics and Design (ISLPED)*, pp.201-206, Aug. 2007.
- [60] E. Kursun and C. Cher, "Temperature Variation Characterization and Thermal Management of Multicore Architectures," *IEEE Micro*, vol. 29, no. 1, pp. 116-126, Jan. 2009.
- [61] V. Gektin, R. Zhang, M. Vogel, G. Xu, and M. Lee, "Substantiation of numerical analysis methodology for CPU package with non-uniform heat dissipation and heat sink with simplified fin modeling," in *Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, pp. 537- 542 Vol. 1, June 2004.
- [62] J. Donald, and M. Martonosi, "Techniques for Multicore Thermal Management: Classification and New Exploration," in *International Symposium on Computer Architecture*, pp.78-88, 2006.
- [63] Y. Ge, P. Malani, and Q Qiu, "Distributed task migration for thermal management in many-core systems," in *ACM/IEEE Design Automation Conference (DAC)*, pp.579-584, June 2010.

- [64] I. Yeo, C. Liu, and E. J. Kim, "Predictive dynamic thermal management for multicore systems," in *ACM/IEEE Design Automation Conference*, pp.734-739, June 2008.
- [65] P. Michaud and Y. Sazeides, "Scheduling issues on thermally constrained processors," Oct. 2006, IRISA report PI-1822 and INRIA report RR-6006; Available: <http://hal.inria.fr/inria-00110085/>.
- [66] K. Skadron, M. Stan, K. Sankaranarayana, W. Huang, S. Velusamy, and D. Tarjan, "Temperature-Aware Microarchitecture: Modeling and Implementation," *ACM Transaction on Architecture and Code Optimization*, Vol. 1, No. 1, Mar. 2004, pp. 94-125.
- [67] M. D. Powell, M. Gomaa, and T. N. Vijaykumar, "Heat-and-Run: Leveraging SMT and CMP to Manage Power Density Through the Operating System," in *International Conference on Architectural Support for Programming Language and Operating System (ASPLOS'04)*, Oct. 2004.
- [68] M. Cho and S. Mukhopadhyay, "Signal processing methods and hardware-structure for on-line characterization of thermal gradients in many-core processors," in *International Symposium on Quality Electronic Design (ISQED)*, pp.797-803, Mar. 2010.
- [69] M. Cho, N. Sathe, M. Gupta, S. Kumar, S. Yalamanchilli, and S. Mukhopadhyay, "Proactive power migration to reduce maximum value and spatiotemporal non-uniformity of on-chip temperature distribution in homogeneous many-core processors," in *Semiconductor Thermal Measurement and Management Symposium*, pp.180-186, Feb. 2010.
- [70] "<http://www.intel.com/cd/channel/reseller/asmo-na/eng/216412.htm>".
- [71] S. Borkar, "Thousand Core Chips – A Technology Perspective," DAC, 2007.
- [72] D. Brooks et. al, "Dynamic thermal management for high-performance microprocessors," HPCA 2001.

- [73] T. R. Conrad et. al, "Impact of moisture/reflow induced delaminations on integrated circuit thermal performance," ECTC, 1994.
- [74] M. Cho et. al, "Optimization of burn-in test for many-core processors through adaptive spatiotemporal power migration," ITC 2011.
- [75] Y. Cheng et. al, "Electrothermal analysis of VLSI system," Kluwer Academic Publishers 2000.
- [76] Y. Zhan et. al, "High-Efficiency Green Function-Based Thermal Simulation Algorithms," IEEE TCAD, 2007.
- [77] R. Cochran et. al, "Spectral Techniques for High-Resolution Thermal Characterization with Limited Sensor Data," DAC 2009.
- [78] ASTM: Standard Test Method for Thermal Transmission Properties of Thermally Conductive Electrical Insulation Materials, Designation D 5470-06, ASTM International, 2006.
- [79] A. Poppe and V. Szekely, "Dynamic Temperature Measurement: Tools Providing a Look into packaging and mount structures", Electronic cooling, 2000.
- [80] Robert Campbell, "Flash diffusivity method: A survey of capabilities", Electronics cooling magazine, 2002.
- [81] S. Y. Kim and R. L. Webb, "Analysis of convective thermal resistance in ducted fan-heat sinks," IEEE Transactions on components, packaging and manufacturing technology, 2006.
- [82] K. Kurabayashi and K. E. Goodson, "Precision measurement and mapping of die-attach thermal resistance," IEEE Transactions on components, packaging and manufacturing technology, 1998.
- [83] Predictive Technology Model (PTM): <http://ptm.asu.edu/>

- [84] Intel 64 and IA-32 Architecture Optimization Reference Manual, Intel Corp. Nov, 2009, pp. 49-61.
- [85] Y. K. Cheng et. al, "An efficient method for hotspot identification in ULSI circuits," IEEE ICCAD, 1999.
- [86] Abdullah Nazma Nowroz, Ryan Cochran, and Sherief Reda. 2010. Thermal monitoring of real processors: techniques for sensor allocation and full characterization. In Proceedings of the 47th Design Automation Conference (DAC '10)
- [87] T. Kemper et. al, "Ultrafast temperature profile calculation in IC chips," International Workshop on Thermal investigations of ICs, 2006.
- [88] D. Schweitzer, "A fast algorithm for thermal transient multisource simulation using interpolated Zth functions," IEEE Transactions on components, packaging and manufacturing technology, Vol. 32, June 2009.
- [89] N. Allec et. al, " ThermalScope: Multi-scale thermal analysis for nanometer-scale integrated circuits ", ICCAD 2008.
- [90] H. Wang et. al, "Composable thermal modeling and characterization for fast temperature estimation," IEEE EPEPS, 2010.
- [91] W. Huang, K. Sankaranarayanan, R. J. Ribando, M. R. Stan, and K. Skadron. "Accurate, Pre-RTL Temperature-Aware Processor Design Using a Parameterized, Geometric Thermal Model" *IEEE Transactions on Computers*, 57(9):1277-88, Sept. 2008.
- [92] P. Zhou et. al., "Thermal effects with leakage power considered in 2D/3D floorplanning," IEEE international conference on computer-aided design and computer graphics, 2007.
- [93] G. Loh, et. al. "Zesto: A cycle-Level Simulator for Highly Detailed Microarchitecture Exploration," ISPASS 2009.

- [94] S. Li, et.al., "McPAT: An Integrated Power, Area, and Timing Modeling Framework for Multicore and Manycore Architecture," IEEE/ACM MICRO, Dec. 2009.
- [95] S. R. Vangal, et. al. An 80-Tile Sub-100-W TeraFLOPS Processor in 65-nm CMOS, IEEE JSSC, vol. 43, no. 1, Jan. 2008, pp. 29-41.
- [96] N. A. Kurd, et. al. "Westmere: A family of 32nm IA processors," ISSCC 2010.
- [97] N. Mehta, et. al, "In-Situ Power Monitoring Scheme and Its Application in Dynamic Voltage and Threshold Scaling for Digital CMOS Integrated Circuits," ISLPED 2010.
- [98] A. Coskun et. al., "Utilizing Predictors for Efficient Thermal Management in Multiprocessor SoCs," IEEE TCAD, 2009.
- [99] I. Yeo et. al., "Predictive Dynamic Thermal Management for Multicore Systems," DAC 2008.
- [100] S. Han et al., "Reverse temperature dependence of circuit performance in high-k/metal-gate technology, IEEE Electron Device Letter, vol. 30, December 2009.
- [101] D. Wolpert, et al. "A Sensor to Detect Normal or Reverse Temperature Dependence in Nanoscale CMOS Circuits," IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems, 2009.
- [102] Y. Tsividis, Operation and Modeling of the MOS Transistor, 1st ed. New York, NY, USA: McGraw-Hill.
- [103] A. Osman et al., "An extended Tanh law MOSFET model for high temperature circuit simulation," IEEE JSSC, 1995.
- [104] K. Wang et al., "The zero-temperature-coefficient point modeling of DTMOS in CMOS integration," IEEE Electron Device Letter, Vol. 31, No 10. pp.1071-1073, Oct. 2010.
- [105] M. Paremeswaran et al., "Micro-machined thermal radiation emitter from a commercial

CMOS process”, IEEE EDL, 1991.

[106] Keithly test script processor manual:

<http://www.keithley.com/products/dcac/currentvoltage/gpmp?mn=2602A>.

[107] U2C-12 SPI controller: <http://www.diolan.com>

[108] W. Song, M. Cho, S. Yalamanchili, S. Mukhopadhyay, and A. Rodrigues, "Simulation Infrastructure for Power and Temperature Modeling in Manycore Processors and Linear System Modeling," *The SRC Premier Technical Conf.*, Sept. 2011.